Universidade do Minho

Escola de Ciências

Tiago Manuel Fontes Oliveira

Exploring the role of *O*-Mannosylation in the modulation of E-Cadherin function in Gastric Cancer cell line models

Tese de Mestrado

Mestrado em Bioquímica Aplicada – Biomedicina

Trabalho efectuado sob a orientação da

Professora Doutora Salomé S. Pinho

Setembro 2015

DECLARAÇÃO

Nome

_____

Endereço electrónico: _____ Telefone: _____ / _____

Número do Bilhete de Identidade: _____

Título dissertação □/tese □

_____

_____

_____

Orientador(es):

_____

_____ Ano de conclusão: _____

Designação do Mestrado ou do Ramo de Conhecimento do Doutoramento:

_____

Nos exemplares das teses de doutoramento ou de mestrado ou de outros trabalhos entregues para prestação de provas públicas nas universidades ou outros estabelecimentos de ensino, e dos quais é obrigatoriamente enviado um exemplar para depósito legal na Biblioteca Nacional e, pelo menos outro para a biblioteca da universidade respectiva, deve constar uma das seguintes declarações:

1. É AUTORIZADA A REPRODUÇÃO INTEGRAL DESTA TESE/TRABALHO APENAS PARA EFEITOS DE INVESTIGAÇÃO, MEDIANTE DECLARAÇÃO ESCRITA DO INTERESSADO, QUE A TAL SE COMPROMETE;

2. É AUTORIZADA A REPRODUÇÃO PARCIAL DESTA TESE/TRABALHO (indicar, caso tal seja necessário, nº máximo de páginas, ilustrações, gráficos, etc.), APENAS PARA EFEITOS DE INVESTIGAÇÃO, MEDIANTE DECLARAÇÃO ESCRITA DO INTERESSADO, QUE A TAL SE COMPROMETE;

3. DE ACORDO COM A LEGISLAÇÃO EM VIGOR, NÃO É PERMITIDA A REPRODUÇÃO DE QUALQUER PARTE DESTA TESE/TRABALHO

Universidade do Minho, ____/____/_____

Assinatura: _____

Agradecimentos

Em primeiro lugar gostaria de agradecer à Professora Salomé e ao Professor Celso Reis por me terem permitido colaborar no grupo "Glycobiology in Cancer", bem como a todos os que me ajudaram a crescer profissional e pessoalmente no instituto. Obrigado a toda a gente do IPATIMUP que esteve ao meu lado a cumprir esta etapa. Vocês são boa gente. Desejo-vos o melhor. Stefan, I cannot let this pass. You are one of the brightest guys I know. I wish you all the best! And maybe we will work together in the future, who knows…

Um obrigado ao Professor Pedro pelo apoio que veio de Braga, sempre disponível.

Para o meu pessoal da Bioquímica, deixo nesta tese a alma do bioquímico FCUP/ICBAS. Viais, Escuta, Márcio, Lau, Melo, Tiago… Vocês são os maiores! Que tenham a maior sorte do mundo, sempre.

Ao meu grande Batista e à sua Tati, obrigado pelos cafés, confidências, cinemas, jantares, copos,… *Valar Morghulis*.

Para a Ana, que me atura todos os dias, a chover, a fazer sol, a rir, a chorar, o maior dos obrigados. Seria incapaz de tudo isto sem ti. A verdade é que existe sempre uma grande mulher à beira de um grande homem, pelo que, a receber algum crédito por este trabalho, terei sempre que o dividir contigo. Ajudaste-me a descobrir quem sou e o que quero para mim. És uma verdadeira *singularidade* na vida de alguém. Tenho a maior sorte do mundo.

À minha família que acredita no que sou e como sou.

À minha avó.

RESUMO

A acção do complexo E-Caderina (E-Cad)-cateninas para a formação de junções aderentes estáveis é essencial para a arquitectura de tecidos epiteliais e para a sua integridade mecânica, contribuindo de forma preponderante para a supressão tumoral. Assim, a disfunção da E-Cad está geralmente associada com a perda das propriedades de adesão das células epiteliais, com a capacidade de invasão, e com metastização. Como a E-Cad é uma glicoproteína, ela pode ser modificada por *N*- ou por *O*-glicanos, especificamente por glicanos *O*-manose (*O*-Man). A *N*-glicosilação e a *O*-manosilação iniciam-se no retículo endoplasmático (ER), seguindo para o complexo de Golgi, e terminando quando a glicoproteína é libertada no citoplasma ou em vesículas para ser transportada para a membrana celular. Os monossacáridos iniciais podem ser estendidos em diversos compartimentos do complexo de Golgi por acção de diversas glicosiltransferases: no caso da *N*-glicosilação, as *N*-acetilglucosaminiltransferases (GnTs) III e V (GnT-III e GnT-V) estão regularmente envolvidas na modificação da estrutura do glicano pela adição de um resíduo de *N*-acetilglucosamina (GlcNAc) bisecting ou β-1,6 GlcNAc branching, respectivamente. Enquanto a formação de *N*-glicanos com bisecting GlcNAc por acção da GnT-III leva à supressão tumoral, elevados níveis de *N*-glicanos β-1,6 GlcNAc branching formados por acção da GnT-V são geralmente associados a tumores altamente metastáticos. Trabalhos prévios do nosso grupo permitiram descrever o impacto da actividade destas duas enzimas na expressão e regulação da E-Cad, especificamente num contexto de cancro gástrico (GC) que não pode ser explicado apenas devido a alterações genéticas e epigenéticas. Recentemente, a *O*-manosilação ganhou um novo interesse quando relacionada com a E-Cad, uma vez que foram descritos diversos locais de *O*-manosilação disponíveis na E-Cad, e que os glicanos *O*-Man presentes na E-Cad contribuem para as suas funções biológicas, nomeadamente para a adesão célula-célula.

Este trabalho pretende compreender o papel da *O*-manosilação na modulação das funções da E-Cad num contexto de desenvolvimento e progressão tumoral, bem como perceber qual a relação, neste contexto, entre a *O*-manosilação e a *N*-glicosilação. Os resultados obtidos indicam claramente que o aumento de diferenciação de uma linha

celular é acompanhado por um aumento no perfil geral de *O*-manosilação na célula, e que a E-Cad apresenta níveis mais elevados de glicanos *O*-Man ligados nesse caso. Por outro lado, observamos que o aumento de glicanos *O*-Man e a diminuição de glicanos GlcNAc branched ligados à E-Cad está relacionado com um fenótipo mais estável.

Estes resultados apoiam o conceito de que a *O*-manosilação da E-Cad é essencial para as suas funções biológicas, e que a ausência desta modificação pós-traducional (PTM) pode ser um dos factores chave para a perda de função desta glicoproteína, o que leva ao desenvolvimento e progressão tumoral. De modo global, o objectivo principal deste projecto é clarificar ainda mais os mecanismos moleculares que conduzem à disfunção da E-Cad no GC e que são fundamentais para o estabelecimento de adenocarcinomas humanos, tendo em vista potenciais novas aplicações na cliníca.

O trabalho aqui apresentado foi desenvolvido como parte integrante de um projecto que visa explorar o perfil de *O*-manosilação da E-Caderina no cancro gástrico, tendo sido desenvolvido por mim e pela Sandra Carvalho (Estudante de Doutoramento que co-orientou este trabalho), culminando com a preparação de um artigo para ser submetido para publicação.

ABSTRACT

The establishment of stable adherens junctions by the action of the E-Cadherin (E-Cad)-catenins complex is essential for epithelial tissue architecture and mechanical integrity, contributing effectively for tumor suppression. Therefore, E-Cad impairment is often associated with loss of adhesive properties of epithelial cells, invasiveness and metastasis. As E-Cad is a glycoprotein, it can be modified by *N*- and *O*-glycans, specifically *O*-mannose (*O*-Man) glycans. *N*-glycosylation and *O*-mannosylation initiate at the endoplasmic reticulum (ER), carry on in the Golgi compartment, and terminate when the glycoprotein is released to the cytoplasm or liberated in vesicles for the cellular membrane. The initial monosaccharide residues can be further extended in the several compartments of the Golgi apparatus by the action of several glycosyltransferases: in the case of *N*-glycosylation, *N*-acetylglucosamyniltransferases (GnTs) III and V (GnT-III and GnT-V) are often involved in the modification of the glycan structure by the addition of a bisecting *N*-acetylglucosamine (GlcNAc) or β-1,6 GlcNAc branching residue, respectively. While the formation of bisecting GlcNAc *N*-glycans by GnT-III has been proven to be tumor suppressor, high levels of β-1,6 GlcNAc branching *N*-glycans formed by the action of GnT-V are usually associated with highly metastatic tumors. Work by our group allowed us to describe the impact of the activity of these two enzymes in E-Cad expression and regulation, specifically in a gastric cancer (GC) context that cannot be explained solely by genetic or epigenetic alterations. Recently, *O*-mannosylation gained interest relating to E-Cad, as it was described that E-Cad presents several available *O*-mannosylation sites, and that the *O*-Man glycans present on E-Cad contribute to its biological functions, namely cell-cell adhesion.

This work aims to understand how *O*-mannosylation modulates E-Cad functions in tumor development and progression, and to figure out the interplay between *O*-mannosylation and *N*-glycosylation in this context. Our results clearly indicate that the gaining of differentiation status of a cell line is accompanied by an increase in the cellular overall *O*-mannosylation profile, and that E-Cad presents higher levels of *O*-Man glycans attached to it in that case. On the other hand, we observed that the

increase in *O*-Man glycans and the decrease of branched GlcNAc *N*-glycans attached to E-Cad is related to a more stable phenotype.

These results support the idea that *O*-mannosylation of E-Cad is essential for its biological functions, and that the absence of this post-translational modification (PTM) may be one of the key elements for the impairment of this glycoprotein, which leads to tumor development and progression. Globally, the main goal of this project is to further clarify the molecular mechanisms behind E-Cad dysfunction in GC that are important in the setting up of human adenocarcinomas, having in mind potential new applications in the clinic.

The work that is present here was developed as part of a comprehensive project that aims to explore the *O*-mannosylation profile of E-Cadherin in Gastric Cancer, carried out by me and Sandra Carvalho (PhD Student that co-supervised this work), which will culminate in the preparation of a manuscript to be submitted for publication.

Table of Contents

Image Index

from digested total cell lysates of MKN28 and KATO III cells with Con A and the antibody against actin (loading control). MKN28 presents higher content in *O*-mannosyl glycans when compared to KATO III, as the lectin blot presents higher reactivity in the MKN28 cell line lane.

Figure 15 - The overall *O*-mannosylation profile of MKN28 and KATO III. Lectin blot and Western blot analysis from digested total cell lysates of MKN28 and KATO III cells with Con A and the antibody against actin (loading control). MKN28 presents higher content in *O*-mannosyl glycans when compared to KATO III, as the lectin blot presents higher reactivity in the MKN28 cell line lane.

Figure 16 - *O*-mannosylation of E-Cadherin. (A) E-Cadherin was immunoprecipitated from MKN28 and KATO III cell lines and the existence of *O*-mannosyl glycans attached was assessed using simultaneously PNGase F digestion and Con A lectin blot. (B) *O*-mannosyl glycans were only detected on E-Cadherin in MKN28 cell line. Level of E-Cadherin was assessed in parallel using the antibody anti-E-Cadherin (clone 36) as a control of the immunoprecipitation. The levels of Con A density for MKN28 and KATO III were obtained after normalization to E-Cadherin.

Figure 17 - Levels of *O*-mannosyl glycans and branched *N*-glycans in total cell lysates of MKN28 and KATO III cell lines. Lectin blots were performed in fresh total cell lysates, and actin was used as a loading control for both of them. An initial step of digestion was necessary when using Con A lectin to assess the *O*-mannosylation status of the lysates.

Figure 18 - E-Cadherin *O*-mannosylation and branched *N*-glycans content in MKN28 and KATO III cell lines. (A) E-Cadherin was immunoprecipitated in order to study its *O*-mannosylation profile and β-1,6 glycans content. Whereas MKN28 E-Cadherin presents a huge level of *O*-mannosyl glycans, it seems that in KATO III E-Cadherin is not very *O*-mannosylated. The levels of L-PHA reactivity in both cell lines are very similar to each other.

Figure 19 - Relative expression of *MGAT5* and *POMT2* mRNA in MKN45 MOCK and MKN45+GnT-V cell lines. (*left*) *MGAT5* is expressed approximately 30 times more in the MKN45 cell line transfected with GnT-V. (*right*) *POMT2* mRNA levels are slightly superior in MKN45+GnT-V when compared to the cells transfected with the empty vector (MKN45 MOCK). Results were normalized for *GAPDH* mRNA expression levels. Relative expression of MKN45 MOCK was considered to be 1, and the expression of MKN45+GnT-V was compared in each condition (Student's t-test: * P≤0.05; **** P≤0.001).

Figure 20 - E-Cadherin *O*-mannosylation and branched *N*-glycans content in MKN45 MOCK and MKN45+GnT-V cell lines. (A) E-Cadherin was immunoprecipitated in order to study its *O*-mannosylation profile and β-1,6 glycans content. The presence of GnT-V (MKN45+GnT-V) increases the number of β-1,6 branched *N*-glycans, when comparing to MKN45 MOCK. The level of reactivity of Con A is higher in MKN45 MOCK cells, which indicate a greater number of *O*-mannosyl glycans present on E-Cadherin, when compared to MKN45+GnT-V. (B and C) The levels of Con A and L-PHA densities for MKN45 MOCK and MKN45+GnT-V were obtained after normalization to E-Cadherin, being expressed as the fold change, compared with the association level of the same lectin with E-cadherin in MKN45 MOCK cells, which was taken as 1 (Student's t-test: **P≤0.01; **** P≤0.001).

Figure 21 – Proposed model for the interplay between *O*-mannosylation and *N*-glycosylation. In a normal epithelial cell context, more *O*-mannosyl glycans are linked to E-Cad, as the protein does not have such bulky *N*-glycans due to the absence of GnT-V activity. On the other hand, in an adenocarcinoma cell context, GnT-V activity leads to the formation of huge complex *N*-glycan structures, which precludes the attachment of so many *O*-Man residues, contributing to a decrease in cell-cell adhesion and to a decreased stability of the E-Cad-catenins complex.

List of Abbreviations

α-DG – Alpha-dystroglycan

aa. - Aminoacid

Asn - Asparagine

β-DG – Beta-dystroglycan

Ca^{2+} - Calcium

C1GalT1 - Core 1 Gal-transferase

C2GnT - Core 2 β1-6 *N*-acetylglucosaminyltransferase

CMDs – Congenital Muscular Dystrophies

Con A – Concanavalin A

CPTM – Co- and post-translational modification

Cys - Cysteine

Dol-P_i – Dolichol phosphate

E-Cad – E-Cadherin

ECM – Extracellular Matrix

ECs – Cadherin domains

E-PHA - Biotinylated *Phaseolus vulgaris* Erythroagglutinin

ER – Endoplasmic Reticulum

Fuc – Fucose

Gal – Galactose

GalNAc – *N*-acetylgalactosamine

GalNAc-T - UDPGalNAc-polypeptide *N*-acetylgalactosaminyltransferase

GC – Gastric Cancer

Glc – Glucose

GlcA – Glucuronic Acid

GlcNAc – *N*-acetylglucosamine

Gln – Glutamine

GnTs – *N*-acetylglucosamyniltransferases

IGF-I -Insulin-like Growth Factor I

IGF-IR – Insulin-like Growth Factor I Receptor

IP – Immunoprecipitation

LOH - Loss of Heterozygosity

L-PHA - Biotinylated *Phaseolus vulgaris* Leucoagglutinin

Man – Mannose

miR - MicroRNAs

N-Cad – N-Cadherin

Neu5Ac – Sialic Acid

ON – Overnight

OST – Oligosaccharyltransferase

PBS – Phosphate-Buffered Saline

PBST – Phosphate-Buffered Saline Tween 20 0,05%

PMT - Protein *O*-mannosyltransferase (in yeast)

PNGase F - Peptide-*N*-glycosidase F

polyLacNAc - Poly-*N*-acetylactosamine

POMGnT1 - Protein *O*-linked-mannose β-1,2-N-acetylglucosaminyltransferase 1

POMGnT2 - Manα1-O- Ser/Thr β1,4GlcNAc-transferase

POMT - Protein *O*-mannosyltransferase (in mammalian organisms)

POMT1 - Protein *O*-mannosyltransferase 1

POMT2 - Protein *O*-mannosyltransferase 2

Pro – Proline

PTM – Post-translational modification

R3A-5a - Rhodanine-3-acetic acid Derivative, Compound 5a

RTK – Receptor Tyrosine-Kinase

Ser – Serine

TAA – Tumor Associated Antigen

TCGA – The Cancer Genome Atlas

Thr – Threonine

WB – Western Blot

WHO – World Health Organization

WT – Wild type

1 INTRODUCTION

1.1 Gastric Cancer

According to 2012 GLOBOCAN assessments, Gastric Cancer (GC) was the fifth most commonly diagnosed cancer-related malignancy worldwide and the third most common cause for cancer-related death in both sexes **(Figure 1)**[1,2]. Despite the fact that the incidence of GC has been declining over the past few decades, an increase in the total number of cases is expected due to the ageing population[3]. Focusing on Portugal statistics, GC takes the fifth place when accounting to overall incidence, and the third when considering the mortality rate[2]. However, GC presents a high incidence to mortality rate when compared to all other types of cancers, being supplanted only by lung cancer.

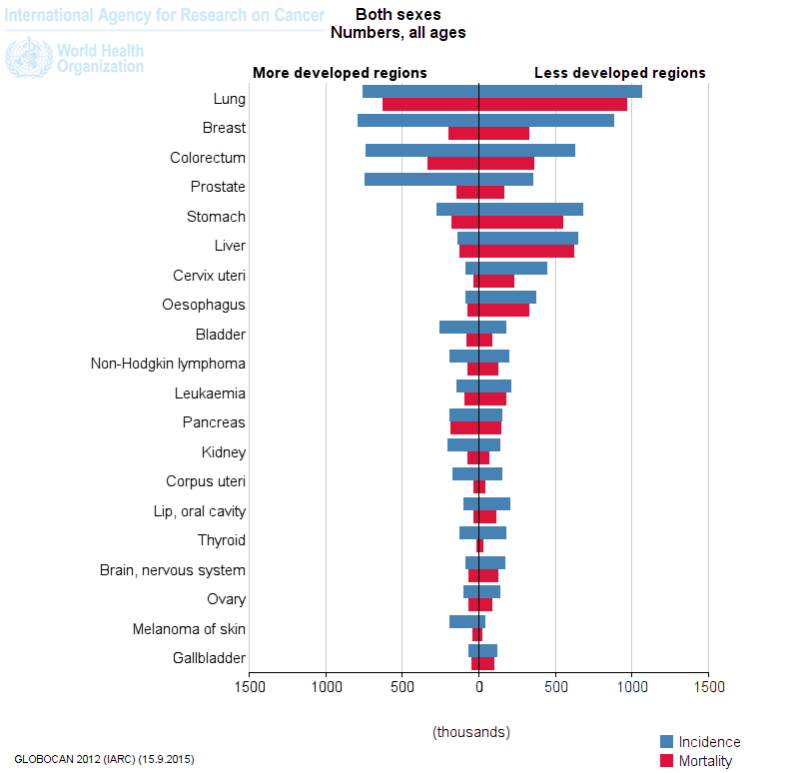

Figure 1 - Overall cancer incidence and mortality estimates worldwide in 2012, according to GLOBOCAN[2].

As GC is a complex and heterogeneous disease displaying different types of phenotypes, several methods were proposed to categorize its specific subtypes, namely the World Health Organization (WHO) and the Laurén classifications[4-6]. In September 2014 The Cancer Genome Atlas (TCGA) Research Network described a whole new way to classify GC into four distinct groups, taking into account its molecular characteristics, specifically: Epstein–Barr virus (EBV)-positive, microsatellite unstable (MSI), genomically stable (GS), and chromosomally unstable (CIN)[7]. However, the Laurén method is still the most described in the large majority of the publications and the most widely used by clinicians.

According to Laurén classification, GC can be sub-divided in two major histological types (according to microscopic morphology of the tumor itself), the Intestinal and the Diffuse types[4]. The Intestinal type comprehends between 60 to 70% of the total number of GC cases[3,6] and is characterized by the formation of autonomous tumors of glandular structure. Several risk factors are associated with this subtype of GC, for example *Helicobacter pylori* (*H. pylori*) infection. *H. pylori* has a central role in different stages of the precancerous processes that culminate in GC, the so-called Correa Cascade[3,8]. Sporadic Intestinal GCs are also usually associated with loss-of-function mutations in the adenomatous polyposis coli (APC) gene and gain-of-function mutations in the *CTNNB1* gene (that codifies β-catenin), which results in an increased signaling via the *Wnt* pathway[9]. In contrast to the Intestinal type, Diffuse GC, which comprehends near 30% of GC cases, is characterized by the loss of cohesion between cells and the invasion of the gastric wall by individually infiltrating neoplastic cells **(Figure 2)**. The development of Diffuse GC and the precise carcinogenic process is still

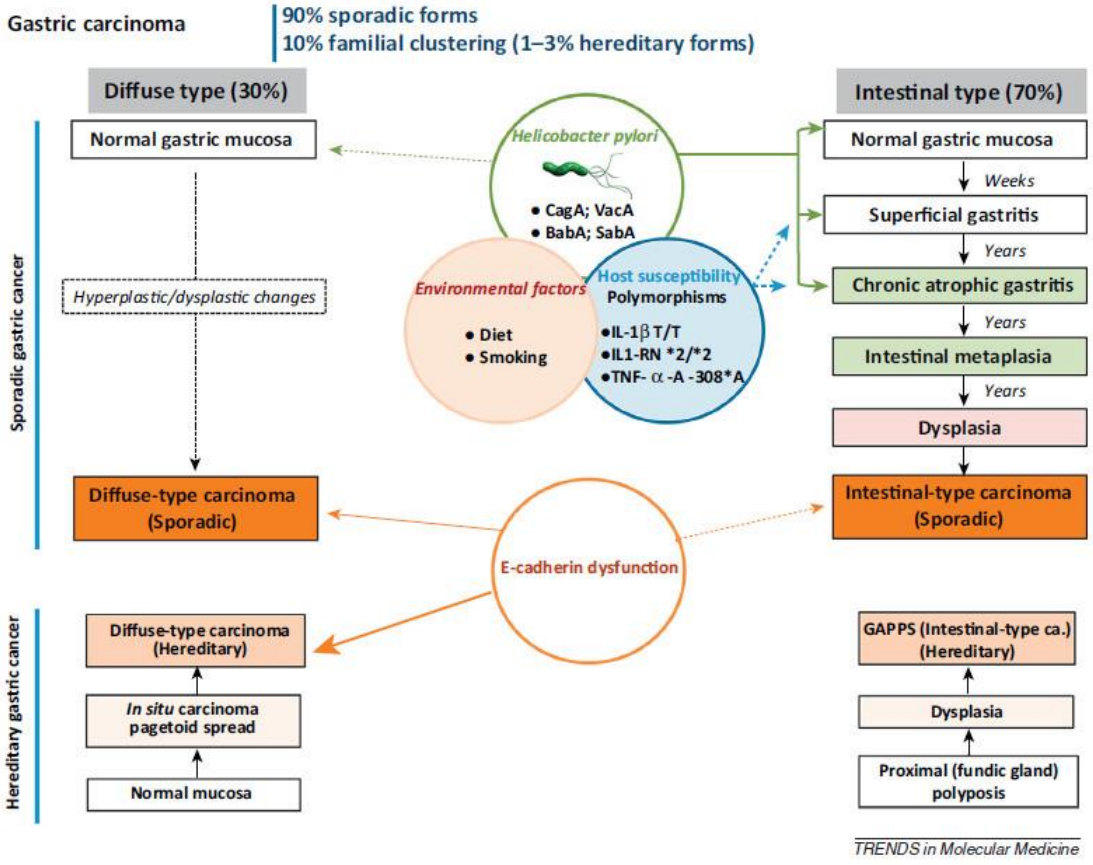

Figure 2 - Clinicopathological profiles of Diffuse and Intestinal Gastric Cancer[11].

debatable, and occurs in the corpus and fundus portions of the stomach[5]. Diffuse GC can also occur due to hereditary conditions associated to genetic abnormalities, namely loss-of-function mutations in the tumor suppressor gene *CDH1*, which encodes for the cell-cell adhesion protein E-Cadherin (E-Cad)[10,11]. As a matter of fact, these mutations in the *CDH1* gene also occur sporadically, which means that loss of E-Cad seems to be a key event in the development of Diffuse GC[12].

Due to their specific phenotypical characteristics, GCs of the Diffuse sub-type are usually diagnosed at latter stages than Intestinal GCs. Usually for both types of GC, surgical resection provides the best approach for effectively remove the tumor. However, late diagnosis presents a problem for the patient, as tumors become very advanced and unresectable. At that time, chemotherapy is the main therapeutic option available, but the five year survival rate drops sharply, with patients surviving usually less than one year. Even though significant efforts are being made in order to improve the therapeutics and surgical approaches, the identification of precise molecular mechanisms and biomarkers is urgently needed to improve early diagnosis, prognosis and for the development of new therapeutic strategies[3,13,14].

1.2 E-Cadherin at the adherens junctions: the importance in cell-cell adhesion

The role of E-Cad in cell-cell adhesion molecule and the importance of calcium (Ca^{2+}) in this process was first described in 1977 by Masatoshi Takeichi[15]. The protein itself was only described in 1980 by François Jacob and his group[16], and it was only one year later, in a subsequent work by the same group, that they named it “uvomorulin”, due to its role in embryogenesis[17]. In 1984, Takeichi *et al*. proposed that the molecules responsible for cell-cell adhesion with the dependence of Ca^{2+}, should be named “cadherins” in a more wide-ranging manner, as an alternative to “uvomorulin”, as they hypothesized that these molecules were present in a diversity of differentiated epithelial cells, instead of being specific of early embryonic cells[18]. Two years later, Takeichi’s group distinguished clearly two individual members of the cadherin family, E-Cad (“E” stands for epithelial) and N-Cad (“N” stands for neural)[19].

E-Cad is a type I cadherin, and is considered to be the prototype of all cadherins due to the fact that it was the first protein from this family that was identified, and to all the efforts that were made for its characterization, which allowed a better understanding of the molecular behavior of E-Cad both in a normal and in a pathological condition. The gene which encodes for human E-Cad, *CDH1*, is located in chromosome 16q22.1, and it was first cloned by Berx *et al.* in 1995[20], which led to subsequent studies regarding the gene activation and silencing (reviewed in [21]). It is interesting to know that Takeichi *et al*. were the first ones to be successful in the attempt to clone E-Cad cDNA (despite it being mouse's cDNA[22]) observing profound alterations on the adhesive properties of the transfected cells. Moreover, the same authors predicted the importance of the E-Cad propeptide (formed by approximately 130 aminoacid (*aa.*) residues) as a short signal sequence for the import of the E-Cad precursor to the ER, as well as the length of E-Cad itself, which was estimated to be 728 *aa.* residues long, comprising a short cytoplasmic domain (151 *aa*.), a single transmembrane segment (26 *aa*.) and a large extracellular domain (551 *aa*.).

The extracellular domain is composed by five cadherin sub-domains (ECs) (EC1 to EC5). These ECs present on cadherins are sequences of 110 *aa*. residues, and are a key element for the classification of this family of proteins. In E-Cad, EC1 is the domain that is responsible for the Ca^{2+}-dependent homophilic interactions between two cadherin molecules of adjacent cells, which is the reason why it is called the "strand swapping domain". The normal conformation of E-Cad is only attained in the presence of Ca^{2+} in the surrounding microenvironment, which is endorsed with the fact that the Ca^{2+}-binding domains are extremely well conserved *aa*. sequences[23]. The other four domains (ECs 2-5) are called the "non-swapping cadherin domains". EC5 is the domain closest to the transmembrane span, and its structure is distinct from the other E-Cad ECs, as it is characterized by the presence of additional four conserved cysteine (Cys) residues, which have an effect in the formation of strong cell-cell contacts[24]. Moreover, E-Cad has four possible *N*-glycosylation sites, two on EC4 (Asn554 and Asn566) and two on EC5 (Asn618 and Asn633)[25], that are of crucial importance when considering E-Cad folding[26] and cell-cell adhesion function[27].

The cytoplasmic domain of E-Cad interacts with different types of catenins, namely α-, β- and p120-catenin. p120-catenin has central role in the maintenance of E-Cad at

the cell membrane, but it also influences the cytoskeleton organization, cell signaling processes and transcriptional regulation[28]. β-catenin is also a central molecule in the cytoskeleton organization, as it is the intermediate molecule between the cytoplasmic domain of E-Cad and α-catenin, which in turn is the responsible for the interaction with the actin cytoskeleton itself[29]**(Figure 3)**.

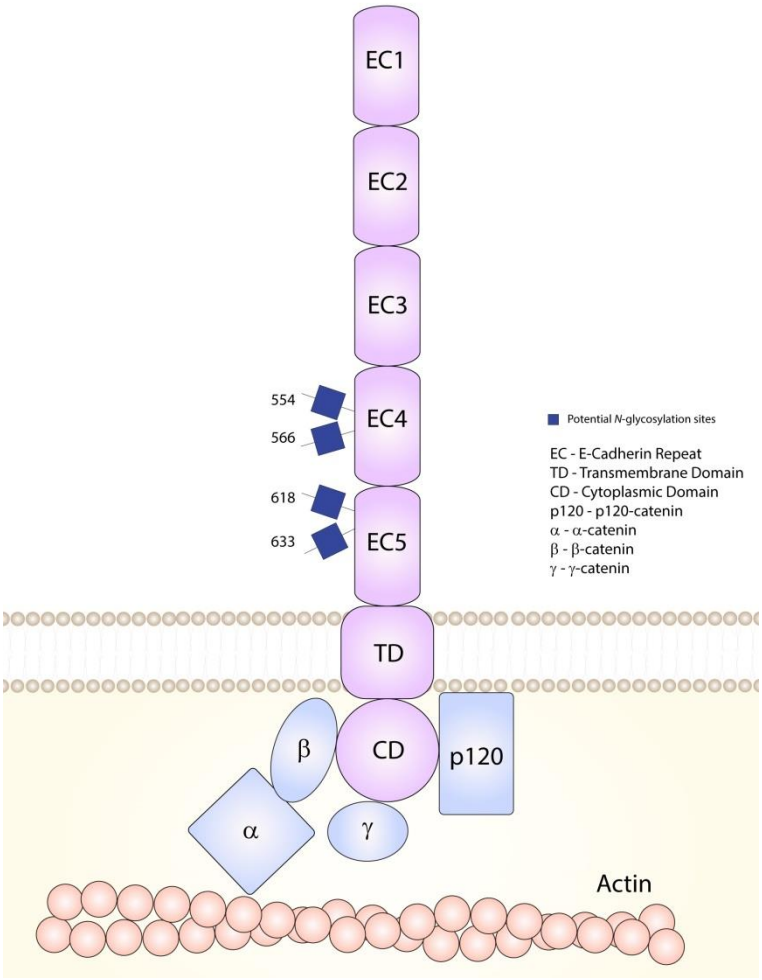

Figure 3 - The E-Cadherin-catenin complex (calcium ions are not represented). Interaction with the actin cytosketelon occurs via α-catenin. The possible *N*-glycosylation sites existent in E-Cadherin (Asn554, Asn566, Asn618 and Asn633) are here represented as the initial GlcNAc monosaccharide.

E-Cad has been extensively described as a key molecule in the maintenance of tissue architecture and homeostasis, by allowing correct epithelial cell-cell adhesion. In order to do so, E-Cad molecules form *cis*- and *trans*- homophilic interactions: t*rans*-homophilic contacts occur when two E-Cad molecules from two different cells establish an adhesive interaction between their EC1s, forming zipper-like structures, while *cis*-interactions are established by E-Cad molecules in the same cell. The *trans*- and *cis*- complexes interact between themselves and with the catenins-actin complex in order to form mechanically stable E-Cad clusters at the adherens junctions, inhibiting cell motility and providing strong and normal associations between cells[29,30].

Several mechanisms have been described to underlie E-Cad dysregulation. Genetic alterations (for example, the loss of exons 8 or 9) of the *CDH1* gene were associated

with absence of a wild type (WT) E-Cad, with decreased cell-cell adhesion and increased cellular motility[31-37]. Epigenetic events, specifically hypermethylation in a CpG island of one 5'-*CDH1* promoter[38], are observable in several types of carcinomas and are have also been associated with loss of E-Cad gene expression. MicroRNAs (miR), such as members of the miR-200 family also regulate *CDH1* expression by targeting transcriptional repressors (in this case, ZEB1 and ZEB2)[39]. All the alterations leading to decreased E-Cad mRNA expression and protein levels are usually associated with loss of cell-cell adhesion and a new capacity of epithelial cells to invade and metastasize, hence a poor clinical outcome [40,41]. The functional loss of E-Cad is one of the best characterized molecular alterations during tumor progression of carcinoma cells (cancer cells arising from epithelial tissues), as emphasized by Hanahan and Weinberg in "Hallmarks of Cancer: The Next Generation"[42]. Moreover, there is a myriad of transcriptional repressors that act frequently on E-Cad promoter (for example, Snail, Slug and Twist[21]), that are specifically expressed at the invasive front of human cancers and are highly regulated by pathways that promote tumor progression, such as Wnt and TGF-β[43]. The protein levels of E-Cad can also be altered due to the action of several receptor tyrosine-kinases (RTKs), such as the endothelial growth factor recetor (EGFR) or the insulin-like growth factor I receptor (IGF-IR), that phosphorylate E-Cad, ultimately resulting in its ubiquitylation by Hakai ligase (and subsequent protein degradation)[43-45].

All these types of alterations can be considered major causes of E-Cad dysfunction in GC[41]. Nevertheless, a significant percentage of GC patients harboring E-Cad dysfunction cannot be explained solely at the genetic or epigenetic levels[46,47]. Several studies have been pointing towards the impact of glycosylation alterations in the regulation of E-Cad in this disease context, as there are several cases that cannot be explained solely by genetic or epigenetic mechanisms[3]. The importance of these post-translational modifications (PTMs) in GC will be further addressed later.

1.3 Protein Glycosylation

Life as we know it cannot be explained solely by the flux of information from DNA to RNA and from RNA to proteins. Proteins can be modified in a wide variety of manners

by a wide variety of molecules, resulting in the huge biological complexity observed. Glycosylation is a universal process of the majority of eukaryotic cells and one of the most frequent occurring PTM, especially when considering transmembrane proteins, since their associated glycans contribute to numerous biological functions, such as cell adhesion, molecular trafficking, signal transduction pathways and endocytosis[48-51]. Glycans exist as the product of covalent attachment of carbohydrate structures, such as oligo- and polysaccharides, to proteins, lipids, carbohydrates or other organic compounds, and can be categorized in several different families of glycoconjugates **(Figure 4)**. The attachment reaction is catalyzed by specific glycosyltransferases that use a specific set of sugars as donor substrates[52]. There are several types of glycans linked to proteins, and the vast majority of those are either *N*- or *O*-linked glycans[50]. *N*-glycosylation is characterized by the addition of oligosaccharides to asparagine(Asn) residues of the proteins, in a consensus sequence Asn-X-serine(Ser)/threonine(Thr), with X being any *aa*. other than proline(Pro)[53,54], while in *O*-glycosylation the attachment reaction occurs in hydroxyl groups of either Ser or Thr residues[55,56].

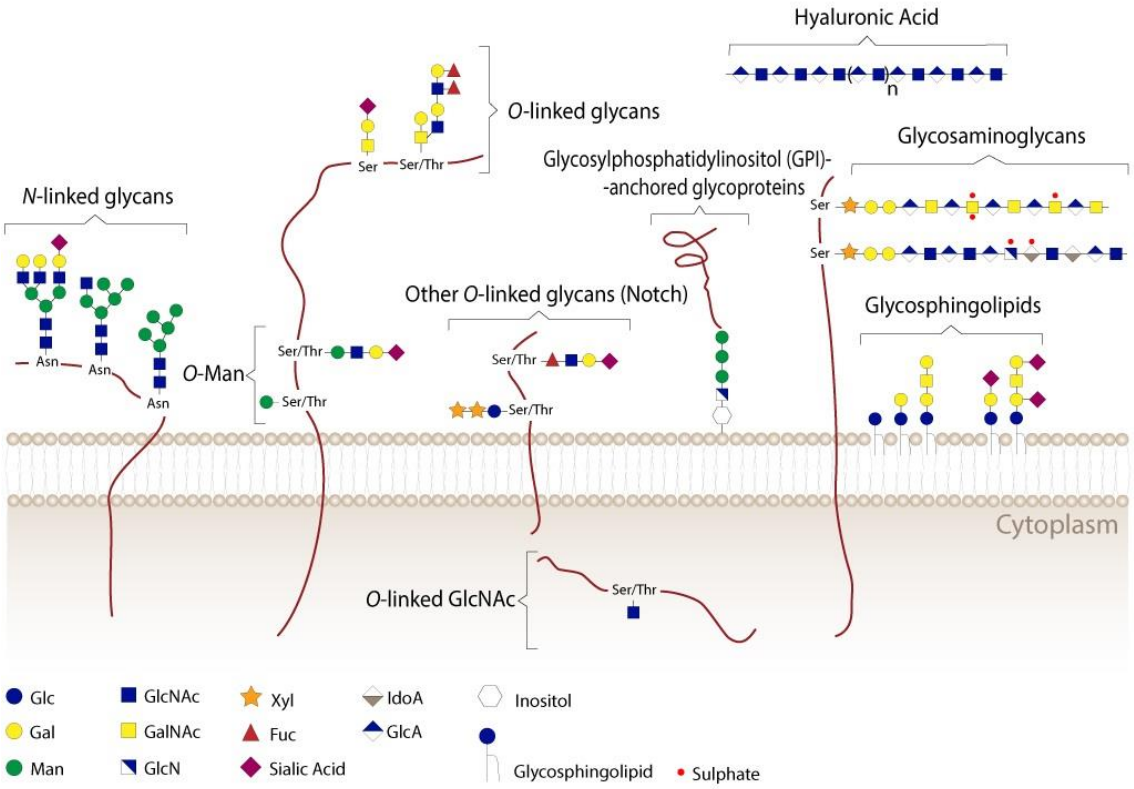

Figure 4 - Major classes of mammalian glycoconjugates.

1.3.1 *N*-glycosylation

N-glycosylation **(Figure 5)** is initiated at the endoplasmic reticulum (ER) membrane where dolichol phosphate (Dol-P_i) located on the cytoplasmic side of the ER functions as a membrane anchor to the formation of the oligosaccharide structure composed by two *N*-acetylglucosamine (GlcNAc) and five mannose (Man) residues, Dol-P_i-P_i-$GlcNAc_2Man_5$. Subsequently, this oligosaccharide structure is flipped to the luminal side of the ER membrane, where four Man and three glucose (Glc) residues are added. This lipid-linked oligosaccharide precursor is then transferred to an Asn residue of a polypeptide chain, that is being translated and arising from the translocon, by the action of oligosaccharyltransferase (OST) enzyme complex[50,57,58], which is why *N*-glycosylation is considered a co- and post-translational modification (CPTM). It is noteworthy that $GlcNAc_2Man_9Glc_3$ acts as a ligand for chaperones from the lectin family such as calnexin and calreticulin, and that the complex formed has a role in protein quality control and folding (the calnexin/calreticulin cycle)[58]. In that process, the Glc residues and one Man residue are removed in the ER by the action of glycosidase II and mannosidase, respectively. The glycoprotein can then continue to the Golgi or be degraded, with its fate depending essentially on its *N*-glycosylation and folding status[59]. In the *cis* compartment of the Golgi apparatus it is common that additional Man residues are removed from the glycoprotein, yielding a $GlcNAc_2Man_5$ glycan linked to the Asn residue. This is the main oligosaccharide structure that allows the sequential action of several *N*-acetylglucosamyniltransferases (GnTs) and glycosidases, resulting in the formation of a complex *N*-glycan. This assembly can then be further modified and extended by the addition of sugar residues such as GlcNAc, galactose (Gal), fucose (Fuc), Glc, and sialic acid(Neu5Ac). There is still a possibility of degradation of badly constructed structures at the *trans*-Golgi, especially the Man-6-P_i-Receptor delivers a largely mannosylated *N*-glycan for the endosomes. Usually however, the final structure is moved from the *trans*-Golgi compartment to vesicles in order to be specifically delivered to the cellular membrane – in the case of transmembrane proteins such as E-Cad – or to be secreted[49,50,54].

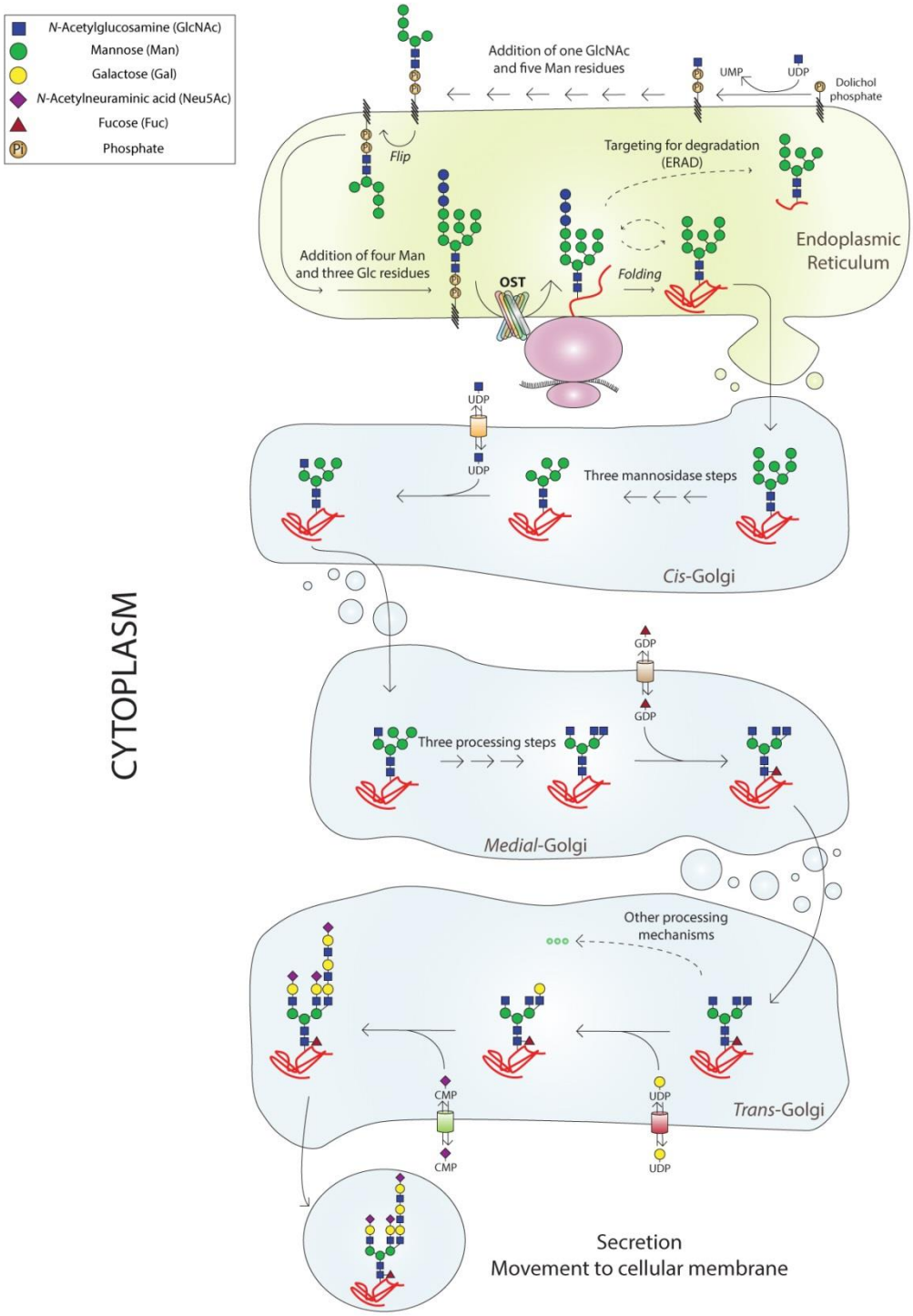

Figure 5 - *N*-glycosylation process occurs in the endoplasmic reticulum and the Golgi compartment. The synthesis of the Dol-P_i-P_i-$GlcNAc_2Man_9Glc_3$ structure is necessary for the action of oligosaccharyltransferase (OST) that transfers the glycan from the activated dolichol carrier to the protein. The protein can then be folded correctly and go to the Golgi compartment, or targeted for degradation. In the Golgi, the action of several glycosyltransferases and glycosidases determines the final glycoprotein that is formed.

1.3.2 O-glycosylation

The most common example of an *O*-glycosylated glycoprotein is the mucin, which has several *N*-acetylgalactosamine (GalNAc) monosaccharides covalently α-linked to several Ser and Thr residues in tandem repeats. This *O*-glycosidic bond formation is controlled by a member of the UDPGalNAc-polypeptide *N*-acetylgalactosaminyltransferase (GalNAc-T) family[60,61]. This GalNAc residue can then be further extended with several monosaccharides such as Gal, GlcNAc, Fuc and Neu5Ac (whereas Man and Glc, for example, are not present in mucin-type *O*-glycans at all). This extension process occurs with the action of several distinct glycosyltransferases (for example, Core 1 Gal-transferase [C1GalT1] and Core 2 β1-6 *N*-acetylglucosaminyltransferase [C2GnT]) and originates eight *O*-GalNAc glycan core structures: cores 1 to 4 are the so-called common cores, and cores 5 to 8 are the additional cores[61,62]. There are several other types of *O*-glycosylation, namely, *O*-mannosylation, *O*-GlcNAcylation, and *O*-fucosylation. For the purpose of the present work, only *O*-mannosylation will be further explored.

1.3.3 O-mannosylation

Protein *O*-mannosylation **(Figure 6)** consists in the covalent linkage of Man to Ser or Thr residues of nascent proteins, and that is the reason why it is considered, as *N*-glycosylation, a CPTM[63]. This process begins in the cytosolic face of the ER with the synthesis of Dol-P_i-Man from Dol-P_i and GDP-Man. The produced Dol-P_i-Man is then flipped to the luminal side of the organelle, where the action of enzymes from the protein *O*-mannosyltransferase family - POMT in mammalian organisms, PMT in yeast (for example) - takes place. These enzymes are responsible for the transference of the Man monosaccharide from the activated lipid donor to the nascent protein, with the inversion of the anomeric configuration of the Man residue added (leading to the formation of an α-D-mannosidic linkage). In mammalian organisms, it seems that the two most important POMTs are POMT1 and POMT2, and that *POMT1* and *POMT2* genes coexpression is necessary in order to achieve the intended enzymatic activity[64]. This *O*-Man can then be elongated in the Golgi apparatus with a GlcNAc

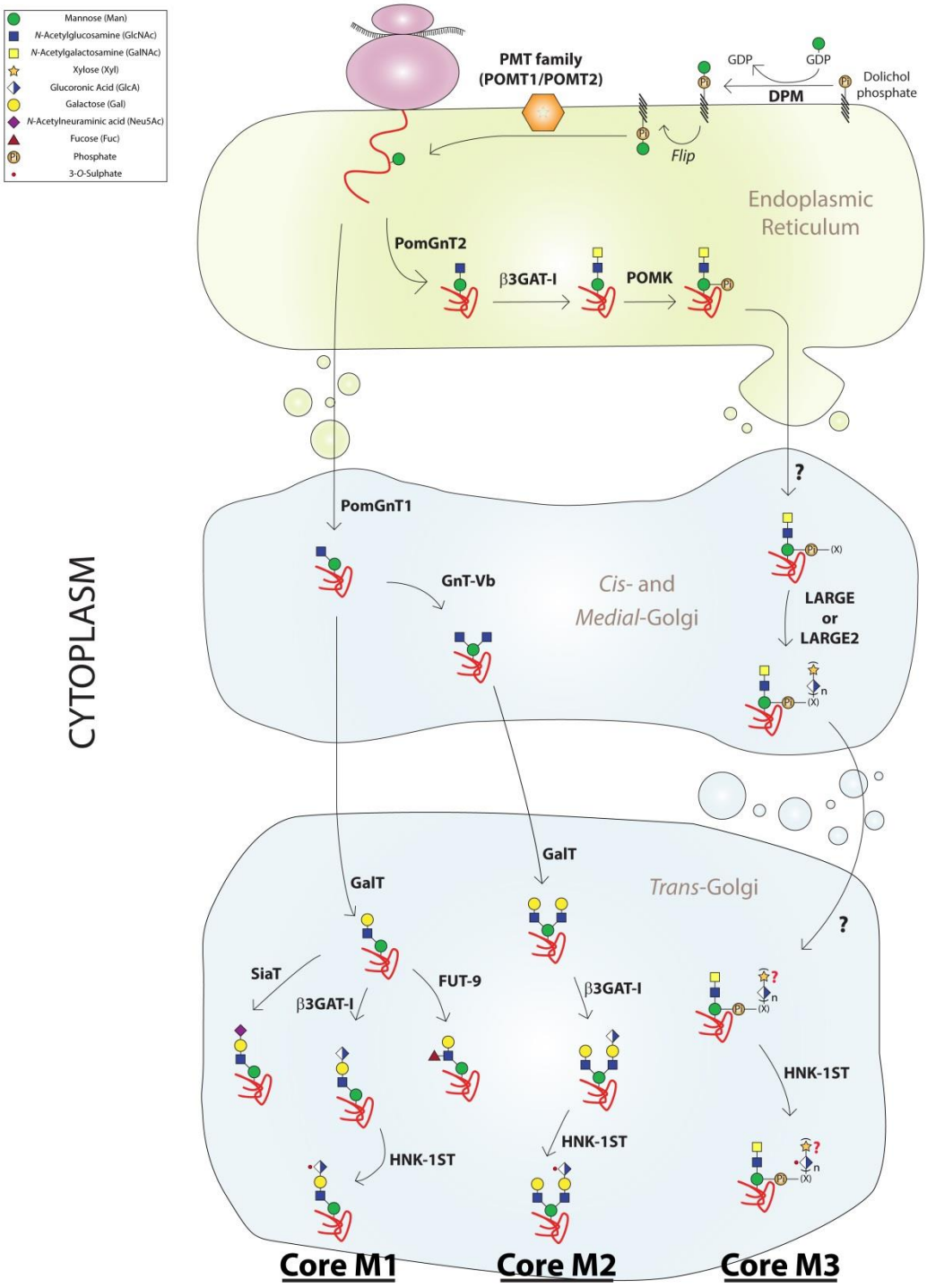

Figure 6 - *O*-mannosylation process initiates in the endoplasmic reticulum by the action of POMT1 or POMT2 that transfer a mannose residue to a nascent protein. The consequent action of several glycosyltransferases determines the type of structure that is formed. *O*-mannosyl glycans can be divided in cores M1, M2 and M3, according to the type of linkage between the mannose and the GlcNAc residues (adapted from [34]).

monosaccharide, in a reaction catalyzed by glycosyltransferases such as protein *O*-linked-mannose β-1,2-N-acetylglucosaminyltransferase 1 (POMGnT1)[65], Manα1-O-Ser/Thr β1,4GlcNAc-transferase (POMGnT2) or GnT-IX (also known as GnT-Vb). These

enzymes do not have, however, a redundant purpose, as the linkage between the GlcNAc and Man sugar residues is not the same, which in turn allows the formation of several *O*-Man-glycans, divided from cores M1 to M3 **(Figure 6)**. This nomenclature to the set of core *O*-Man structures was first proposed in 2013 in a publication from the Campbell laboratory[66], and revised by Jeremy Praissman and Lance Wells in 2014[67], in order to approximate to the one that is accepted for *O*-GalNac glycans. It is also worth mentioning that not all these enzymes have been observed in every tissue studied, which can possibly imply a tissue-specific *O*-mannosylation pattern[67].

This core structures can then be extended by the addition of a variety of monosaccharides, such as Gal, Neu5Ac, Fuc or glucuronic acid (GlcA). At least 23 different *O*-Man glycan structures have been described by now, the majority of which were detected in mammalian brain tissue[67-70]. The constant improvements of glycomic and glycoproteomic technologies will be of huge impact in the attempt to unravel novel *O*-Man structures that can, for instance, be used as important biomarkers.

1.4 Glycosylation in gastric cancer: the importance of *N*-glycosylation of E-Cadherin in a gastric cancer context

Glycans have a particular significance in cell biology when considering the cell-extracellular matrix (ECM) interactions, cell-cell adhesion, cell signaling and communication, and in the immune system **(Figure 7)**[51,71]. It is well known that aberrant glycosylation occurs in all types of cancer, and that there are many specific glycosyl epitopes that constitute the tumor associated antigens (TAA), used often as biomarkers for cancer detection[72-74]. Alterations in the *N*-glycosylation pattern of cell-cell adhesion molecules such as cadherins[3,25,27,75,76] have been correlated with cancer invasion and metastasis.

The frequent dysfunction of E-Cad expression observed in Diffuse-type GC patients was observed by Joo *et al*. in 2002[46]. In effect, this decreased expression was found at earlier stages of Diffuse GC, while in the Intestinal subtype the alteration occurs typically at later stages. Moreover, when the alteration on E-Cad expression is observed in any sub-type of GC, it is usually associated with a poor prognosis for the

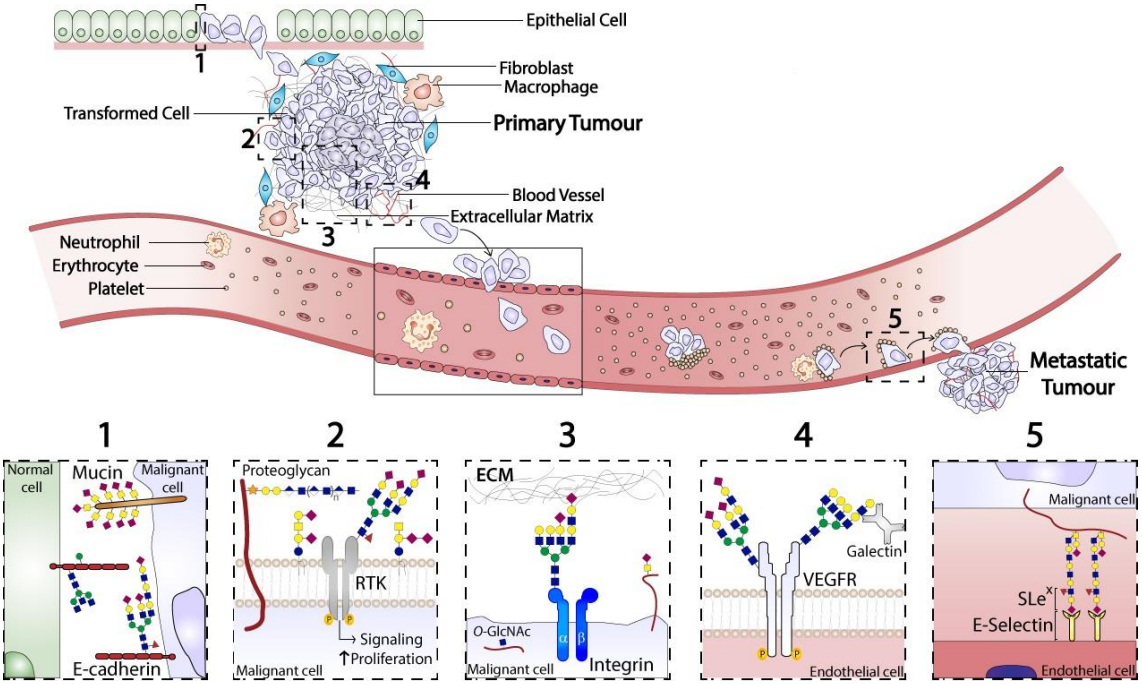

Figure 7 - The importance of glycosylation in the formation of a primary tumor and in the metastatic process. Glycans are important in the cell-cell adhesion process, in cell-extracellular matrix interactions, as well as in several other processes.

patient[3]. It has been described that approximately 70% of GC cases present any structural alteration in *CDH1* gene, but rather a delocalization of E-Cad from the cell membrane to the cytoplasm[47], which suggests that other mechanisms rather than genetic or epigenetic are regulating E-Cad functions in cancer[77,78].

It is now widely accepted the fundamental role of glycosylation (particularly *N*-glycosylation) in the correct molecular organization of E-Cad in the adherens junctions[76], as well as its impact in the process of E-Cad expression, and in the modulation of its biological functions[27,79]. As previously referred, E-Cad has four potential *N*-glycosylation sites, two in EC4 and two in EC5. The *N*-glycans that are attached to E-Cad depend on an intricate network of enzymes (glycosyltransferases) giving rise to different glycans depending on the tissue type and the physiopathological context. The effect of the different glycans on E-Cad activity has been demonstrated, for example, by the enzymatic competition between the *N*-acetylglucosaminyltransferase-III and *N*-acetylglucosaminyltransferase-V (GnT-III/GnT-V) glycosyltransferases. GnT-III is encoded by the *MGAT3* gene, and is responsible for the addition of a GlcNAc residue via β-1,4 linkage to the central Man of the [Asn]-$GlcNAc_2Man_3(GlcNAc_2)$ core structure, while GnT-V, encoded by the *MGAT5* gene, adds

GlcNAc via β-1,6 linkage to an external α6-Man[80]. GnT-III is considered of vital importance in *N*-glycan biosynthesis, as its action precedes that of GnT-V. Therefore, the bisected oligosaccharide cannot be used as a substrate by GnT-V, precluding the synthesis of branching structures[81,82].

In highly metastatic tumors, it is common to observe an increased activity of GnT-V and, consequently, an increased expression of β-1,6 GlcNAc branching *N*-glycan structures[83]. On the other hand, when GnT-V is not present (i.e. *MGAT5-knockout* mice) cancer metastasis is suppressed in a noticeable manner[84]. GnT-III has been described to exhibit an enzymatic priority, which contributes to the suppression of cancer metastasis by increasing the number of bisecting GlcNAc structures and, at the same time, by impeding the formation of branching GlcNAc structures[85]. In addition, an improvement in the cell-cell adhesion was documented when GnT-III was being expressed, as a result of an E-Cad arrest at the cell membrane and delayed turnover[75]. In 2008, the influence of the E-Cad-catenin complex on GnT-III expression was determined, suggesting that this complex may act as an inductor of GnT-III expression, which in turn allows the formation of bisected GlcNAc *N*-glycans[77]. The relationship between E-Cad overexpression and the inhibition of the RTKs IGF-IR and the insulin receptor (IR) was described by our group using MDA-MD-435 cancer cell line (that endogenously does not express E-Cad at the mRNA or protein levels)[86], where stimulation of these cells with insulin or insulin-like growth factor I (IGF-I) induced a decrease in the quantity of bisecting GlcNAc *N*-glycans (in general and attached to E-Cad), that was accompanied by delocalization of E-Cad to the cytoplasm and with an alteration to a more fibroblastoid-like appearance[86]. Furthermore, Pierce *et al*. determined that the action of GnT-V on N-Cad's EC2 and EC3 contributed to a decrease of homotopic cell-cell interactions and to the decrease of several intracellular (outside-in signaling) pathways also involved in cell adhesion (levels of phosphorylated-ERK are enhanced when GnT-V was knockdown)[87].

Taking into consideration that E-Cad dysfunction has a central role in GC development and progression, our group has been particularly interested in characterizing the role of glycosylation as a key molecular mechanism responsible for E-Cad impairment in cancer. In 2009, the group established a clear association between E-Cad expression and the upregulation of the GnT-III transcription levels.

Furthermore, the decrease of β-1,6 branched structures (catalyzed by GnT-V) was noticeable when a lectin specific for this type of structures was used (Biotinylated *Phaseolus vulgaris* Leucoagglutinin [L-PHA]), together with a slight increase of bisecting GlcNAc structures (catalyzed by GnT-III). The role of GnT-III in the regulation of E-Cad transcription was also studied: using siRNA to silence the gene that codifies for the enzyme, the group observed that the level of E-Cad's mRNA remained unaltered, when compared to control cells (non-silenced). Nevertheless, silencing of GnT-III induced a significant alteration in the localization of the protein, with E-Cad exhibiting an increased cytoplasmic expression, and mislocalization from the membrane. Interestingly, the decrease in bisecting GlcNAc structures on E-Cad, detected by *biotinylated Phaseolus vulgaris Erythroagglutinin* (E-PHA - lectin that specifically recognize bisecting GlcNAc *N*-glycans), together with an increase of β-1,6 branched glycans, was observed when GnT-III was silenced. *In vivo* studies were performed in diffuse gastric carcinomas, and a significant mislocalization of E-Cad expression was observed in tumors, concomitantly with a decrease in bisecting GlcNAc structures and higher levels of the branching structures, which validated the *in vitro* results[78]. Moreover, our group also described the impact of GnT-V activity in E-Cad expression in a GC context, and how the absence of this glycosyltransferase (and its product) results in a normal distribution of E-Cad in the cell. Also, the importance of specific *N*-glycosylation sites was explored, with Asn554 as the selected site that is modified with β-1,6 GlcNAc *N*-glycans structures, thereby assuming a central role in the regulation of E-Cad biological functions, namely by interfering with the formation of the E-Cad-catenins complex[27].

Considering all these results, it seems that there is a mutual regulatory mechanism between GnT-III expression and E-Cad-mediated cell-cell adhesion[80], and that the modification of E-Cad by bisecting GlcNAc *N*-glycans induces the recruitment of catenins to the adherens junction, stabilizing the adhesion complex, and impairing cell invasion and metastasis **(Figure 8)**. Accordingly, GnT-V activity in a GC background leads to the formation of tri- or tetra-antennary *N*-glycans that can be extended in a wide variety of ways, such as poly-*N*-acetylactosamine (polyLacNAc) - that is composed by a polymer of GlcNAc and Gal monosaccharides - or terminal Neu5Ac residues[88]. The modification of E-Cad with β-1,6 GlcNAc branched structures affects cell-cell adhesion

capacity[27], by altering the cellular localization of E-Cad and by modifying the cell morphology.

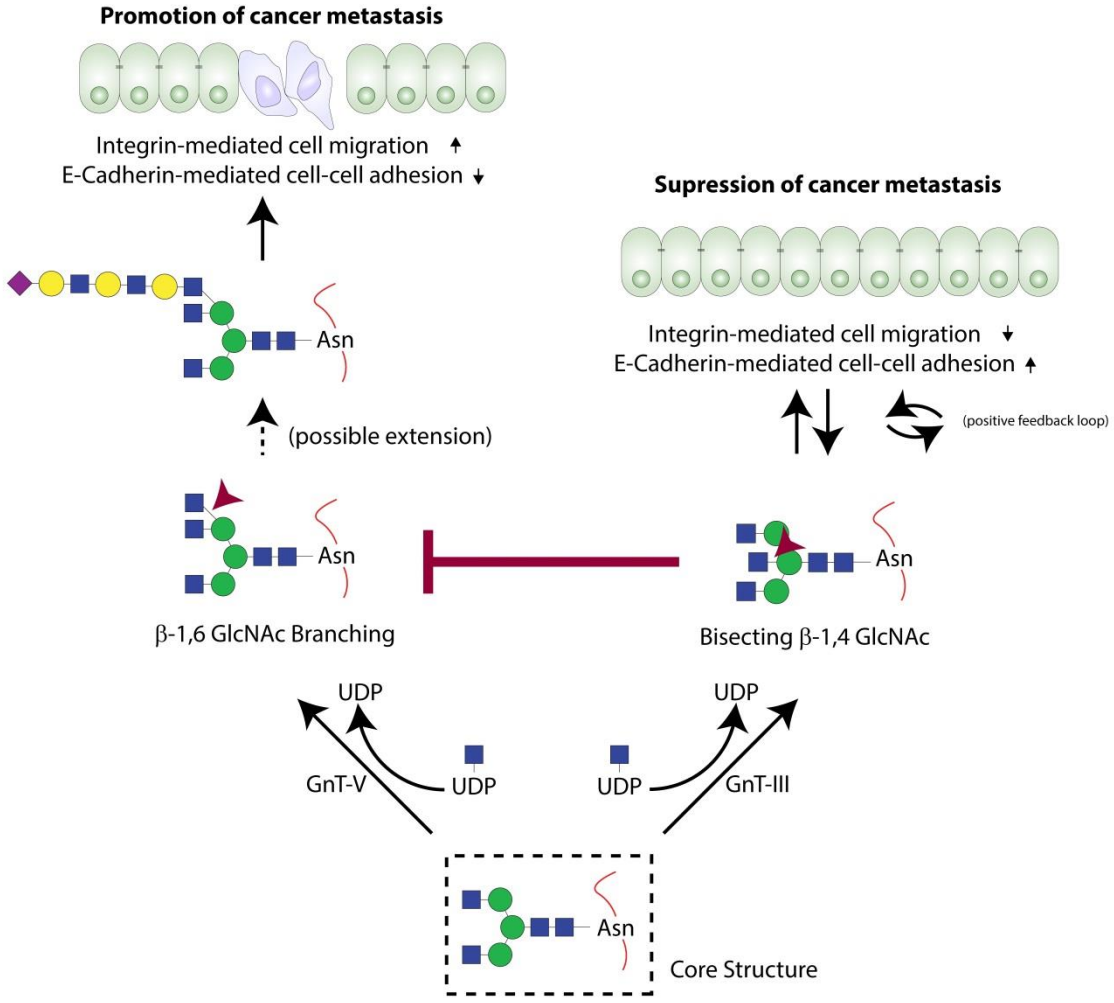

Figure 8 - Alterations in the structure and function of E-Cadherin occur due to the action of GnT-III and GnT-V. Enzymatic activity of these glycosyltransferases determines the type of glycans formed, which in turn contributes to the promotion or suppression of cancer metastasis (adapted from [67]).

1.5 O-mannosylation and E-Cadherin: evidences of a "new" player in the adhesion game

Glycoproteins can be modified in a comprehensive variety of ways, such as the addition of *O*-Man monosaccharides to its Ser/Thr residues contributing to their stability and biological functions. This *O*-mannosylation process was thought to be for a long time, exclusive of organisms such as *Saccharomyces cerevisiae* (*S. cerevisiae*). The first time *O*-mannosylation was described was in 1968 by Sentandreu and Northcote, in a glycopeptide that composed the yeast cell wall[89]. It was only in 1993

that Strahl-Bolsinger *et al*. described the gene *PMT1*, which codifies for the protein PMT1 responsible for the initial step of the yeast *O*-mannosylation process[90]. The same group described a few years later the existence of five additional *PMT* genes in *S. cerevisiae*, which led them to hypothesize the existence of different protein substrate specificities to the different PMTs[91].

In 1996, however, the idea that *O*-mannosylation was a PTM exclusive from yeast was abandoned, as Martín-Blanco and García-Bellido proved the existence of a protein in *Drosophila melanogaster* (*D. melanogaster)*, *rotated abdomen* (*rt*), that presented a high homology to the *S. cerevisiae* PMTs[92]. It is interesting that in *Drosophila* and in *Saccharomyces*, both *rt* and PMT, respectively, exhibit important roles in the development of the organisms, and that mutations in those genes were potentially lethal to the organisms carrying them[93].

In 1999, Jurado *et al*. identified a putative Human homolog of *D*. *melanogaster*'s *rt*, which they named Protein *O*-mannosyltransferase 1 (POMT1). At that time, they were capable of assigning the gene *POMT1* to chromosome 9q34.1, and computer simulations allowed them to predict that the possible product POMT1 would be a transmembrane ER protein[94]. At that moment, and bearing in mind what was known in *Drosophila rt*, the authors suggested that *POMT1* gene mutations might be the explanation for some uncharacterized forms of congenital muscular dystrophies. Later, the first review of the *O*-mannosylation process in mammalian organisms was published, and the molecular organization of the dystroglycan complex, which is composed by α-dystroglycan (α-DG), β-dystroglycan (β-DG), laminin, dystrophin and actin, was illustrated for the first time (this complex is of great importance, in a biological context, as it is responsible for connecting the ECM and the cellular cytoskeleton in skeletal muscle and in Schwann cells). Tamao Endo proposed that *O*-mannosylation should has an important role in the basic biological function of α-DG (the protein linker between laminin and β-DG) as this glycoprotein was depicted as being heavily *O*-mannosylated[56]. A lot of work has been done since, especially relating *O*-mannosylation and α-DG in a pathological context, and it is now known that defects in this process result in a heterogeneous group of congenital muscular dystrophies (CMDs)[66,69,95,96]. The most severe type of CMD is called the Walker-Warburg syndrome, and is essentially characterized by malformations of the brain of the subject, often

resulting in a fatal outcome[97]. The study of CMDs extensively contributed to the understanding of the *O*-mannosylation process. Nevertheless, in 2013 new insights about the *O*-mannosylation process were provided, but this time referring to the association between *O*-mannosylation and cadherins.

Vester-Christensen *et al*. modified a cell line (MDA-MB231, which is derived from a metastatic site of mammary gland, but has an epithelial morphology) to solely exhibit *O*-Man in its *O*-glycoproteome[98]. The fact that the cells were unable to produce extended *O*-GalNAc glycans due to the *knockout* of *COSMC* (the chaperone necessary for C1GalT1 activity), or to further extend the *O*-Man monosaccharide due to the *knockout* of *POMGNT1*, allowed them to look exclusively to the *O*-mannosylation sites existent in all kinds of proteins present in the cell. In order to do that however, they first needed to get rid of the *N*-glycans and the short GalNAc *O*-glycans present in the cell samples, which they did by performing two enrichment steps, using peptide-*N*-glycosidase F (PNGase F - to remove the *N*-glycans) at first, and subsequently using a short Concanavalin A (Con A) column (Con A binds α-Man residues, and the *O*-GalNAc glycans are eluted). The analysis of the nanoflow liquid chromatography-mass spectrometry results confirmed the existence of a large number of *O*-Man glycosites on α-DG, as it was expected. But the most intriguing fact is, of a total of 52 glycoproteins identified, 37 of those belonged to the cadherin family of proteins **(Figure 9)**, which accounted for approximately 56% of the total number of *O*-Man

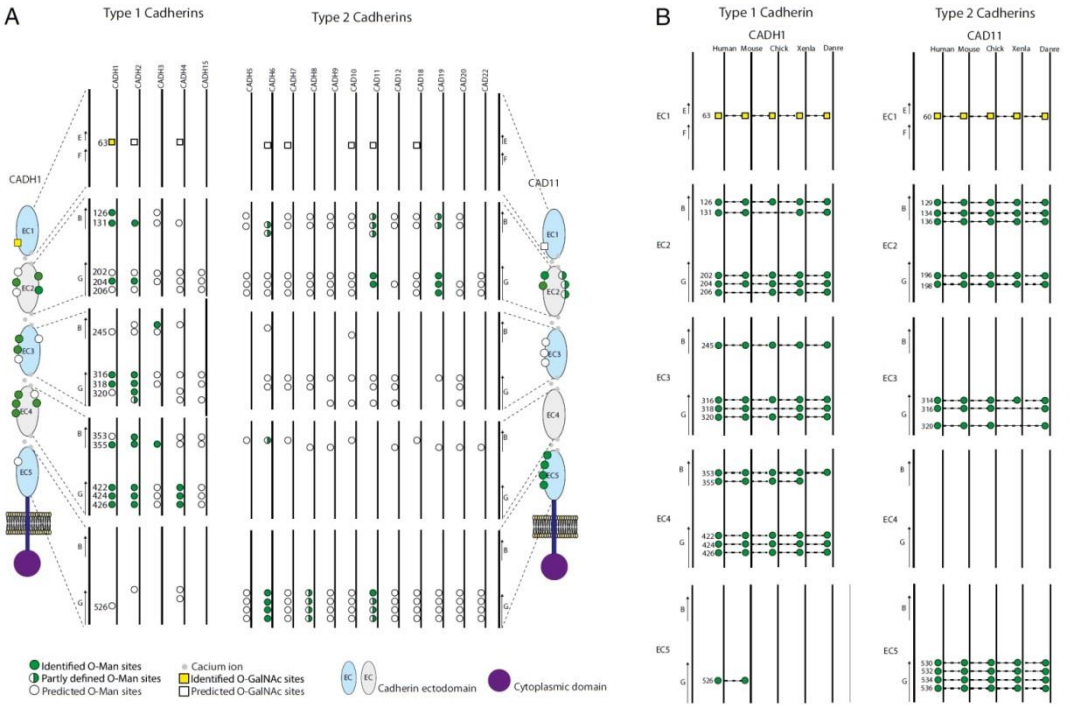

Figure 9 - Type 1 cadherins were identified by the Clausen group as major carriers of *O*-mannosyl glycans[84].

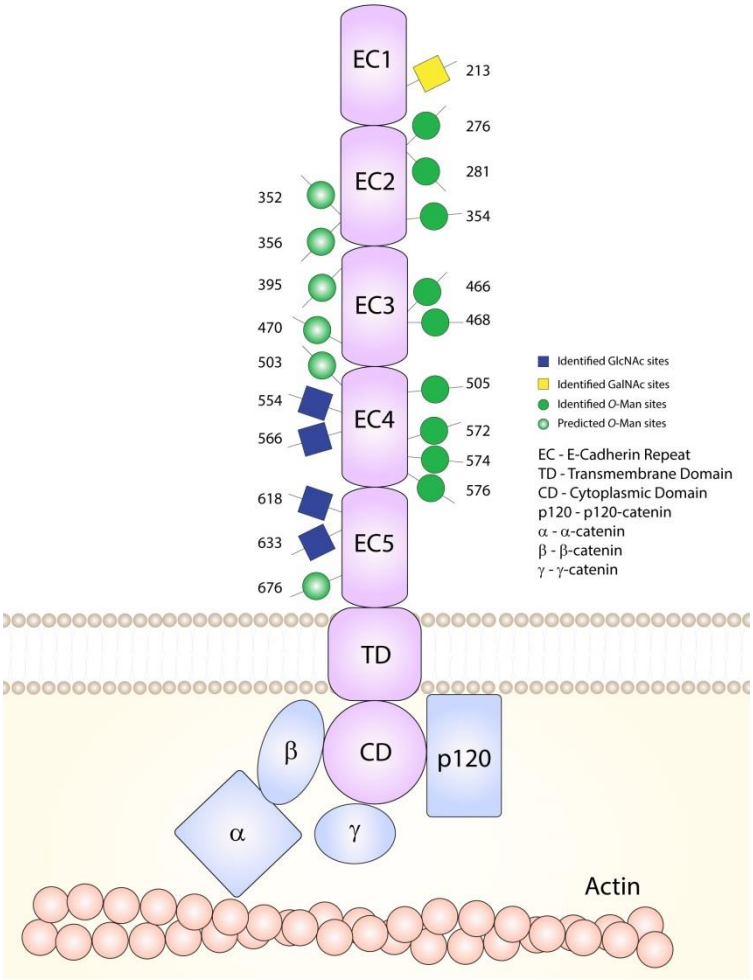

Figure 10 - Glycosites identified and predicted on E-Cadherin to date. There are a total of 4 *N*-glycosylation sites, 9 *O*-mannosylation sites (identified in [84]) and 1 *O*-GalNAc site, distributed on the protein as depicted. A total of 6 *O*-mannosylation sites were also predicted by the Clausen's group. The glycosylation sites represented were numbered considering the propeptide as part of E-Cadherin protein.

glycosites covered on those 52 glycoproteins. On E-Cad, a type 1 cadherin, nine *O*-mannosylation sites were identified, distributed between EC2 (3), EC3 (2) and EC4 (4), and six sites were assigned as putative on ECs 2 to 5 (**Figure 10)**.

Simultaneously, Lommel *et al*. explored the importance of *O*-mannosylation on E-Cad's biological function[99]. In order to do that they used *Pomt2* deficient/null mice, as the adhesion process (and E-Cad in particular) has a central role in the epithelial to mesenchymal transition during embryogenesis (in a previous study by the same group they demonstrated that the absence of *Pomt1* is lethal to mouse embryos[100]). *Pomt2*$^{+/+}$ WT and heterozygous *Pomt2*$^{+/-}$ mice developed normally, and in a very similar manner. However, intercrosses between *Pomt2*$^{+/-}$ mice did not generate offspring, as the *Pomt2*$^{-/-}$ inheritance is lethal even at the embryo phase. The authors isolated these embryos and cultivated them *in vitro*, which allowed them to perceive the inability of *Pomt2*$^{-/-}$ to generate the blastocyst. Additionally, the authors used an inhibitor of fungal POMTs rhodanine-3-acetic acid derivative, compound 5a (R3A-5a), to selectively inhibit the activity of POMT in Madin-Darby canine kidney (MDCK) cells. They observed using an antibody (VIA4-1) specific for an *O*-Man linked epitope on α-DG, that 50 μM of R3A-5a was enough to inhibit the *O*-

mannosylation of α-DG. Having this in mind, the group used R3A-5a when cultivating embryos, which arrest their development at the morula to blastocyst transition, with the phenotype observed being in everything similar to that of the *Pomt2*$^{-/-}$ embryos. To validate all these results, the authors generated an antibody, T[α1-Man], against a Thr *O*-Man-conjugated epitope, to detect *O*-mannosylated proteins. The conjugated use of T[α1-Man] with the inhibitor R3A-5a and *Pomt2*$^{-/-}$ mice allowed them to better understand the importance of *O*-mannosylation in cell-cell adhesion, and how that influences the development of the embryo. It seemed that, in the R3A-5a-treated embryos, the failure in the morula-to-blastocyst transition was due to an abnormal formation of cell-cell contacts, namely adherens junctions, which suggested that defects in the *O*-mannosylation of E-Cad could lead to an impairment of cell-cell adhesion. In fact, aggregation assays performed using MDCK cells (that express high endogenous levels of E-Cad) suggest that *O*-mannosyl glycans affect cadherin mediated cell adhesion, specifically due to the high aggregation inhibition observed when using the T[α1-Man] antibody and an E-Cad blocking antibody **(Figure 11)**.

The influence of glycosylation for E-Cad function has been extensively described, namely the influence of two specific glycosyltransferases – GnT-III and GnT-V – in the *N*-glycan landscape composing this transmembrane protein. Considering this relationship, and taking into account the evidences that indicate that E-Cad *O*-mannosylation is essential for its correct biological function, the main purpose of this project is to assess how *O*-mannosylation, and its interplay with *N*-glycosylation, regulate E-Cad adhesion function, and how this process ultimately influences cancer development and progression.

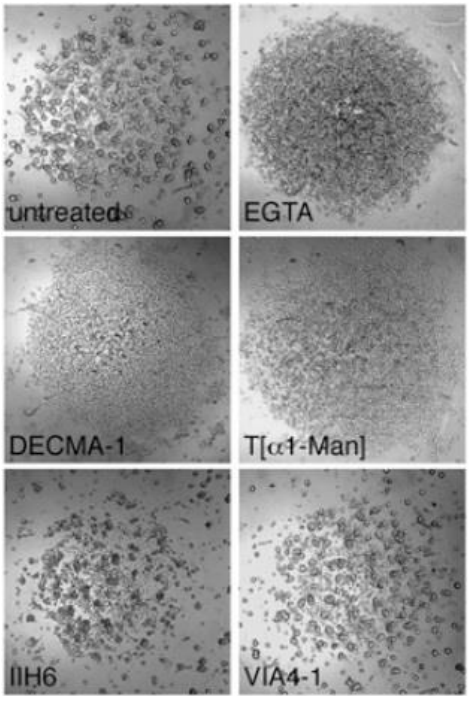

Figure 11 - Slow aggregation assays performed using the MDCK cell line using an E-Cadherin antibody (DECMA-1) and a T[α1-Man]-specific antibody allowed to understand the importance of *O*-mannosyl glycans in cell-cell adhesion. EGTA (blocks the Ca^{2+}-dependent cadherin-mediated adhesion), IIH6 and VIA4-1 (antibodies directed against *O*-mannosyl epitopes linked to α-dystroglycan) were used as controls[86]

2 AIMS

Aims

The main purpose of this project is to understand the role of *O*-mannosylation and its interplay with *N*-glycosylation in the modulation of E-Cad functions in tumor development and progression.

In order to accomplish this goal, three specific aims were designed:

1) To evaluate the *O*-mannosylation profile in different GC cell line models

2) To determine the pattern of E-Cad *O*-mannosylation in a cancer context using *in vitro* cellular models, in order to evaluate the importance of this PTM on E-Cad biological function

3) To characterize the relationship between *N*-glycosylation and *O*-mannosylation of E-Cad in cancer

3 MATERIALS AND METHODS

3.1 Cell lines and cell culture

MKN28 (JCRB0253) and KATO III (ATCC HTB-103) cell lines were cultured at 37ºC in an incubator with 5% CO_2 in RPMI 1640 medium with Glutamax (Gibco, Invitrogen) containing 10% of fetal bovine serum (FBS) (Gibco, Invitrogen), and 1% of the antibiotics penicillin/streptomycin (P/S; Invitrogen).

MKN45 MOCK (empty vector) and MKN45+GnT-V (vector containing *N*-acetylglucosaminyltransferase V) cell lines had already been previously stably transfected[81,101], and were cultured at 37ºC in an incubator with 5% CO_2 in RPMI 1640 medium with Glutamax (Gibco, Invitrogen) containing 10% of FBS (Gibco, Invitrogen), 1% of the antibiotics P/S (Invitrogen), and 500 μg/mL of the selection antibiotic G418 (Gibco, Invitrogen).

3.2 Immunofluorescence

Cells from MKN28 and KATO III cell lines were cultured in coverslips in 6 well plates. When confluent, the medium was removed and the cells were washed 3 times with PBS. After that, the 6 well plates were left on ice as the cells were fixed with methanol (previously stored at -20ºC). The cells in the coverslips were then washed 3 times with PBS, and subsequently blocked with a BSA 5% solution in PBS for 30 minutes at room temperature. After that, the primary mouse monoclonal anti-E-Cad IgG2a antibody (clone 36, BD Transduction Laboratories and Cell Signaling), diluted 1:200 in the BSA 5% solution in PBS, was used to incubate the coverslips for 1 hour at room temperature. The coverslips were then washed 3 times with PBS and, from this point forward, the procedure was made by avoiding light contact with the samples. Incubation with *goat Alexa Fluor 488 anti-mouse* (1:500 – Invitrogen) was followed by 6 consecutive washing steps with PBS and by the use of DAPI (1:100) as a nuclear marker. A VECTASHIELD *mounting medium for immunofluorescence* kit was used, and the slides were stored at -20ºC and in the dark, until visualization in a Carl Zeiss Apotome Axiovert 200 M Fluorescence Microscope.

3.3 PNGase F digestion

PNGase F digestion was performed in order to analyze *O*-linked α-Man glycans specifically, avoiding the interaction of the α-Man residues present in the *N*-linked glycans with the lectin used (Con A). PNGase F is an amidase that cleaves that cleaves between the innermost GlcNAc and Asn residues of high-Man, hybrid and complex oligosaccharides from *N*-linked glycans in proteins. Total cell lysates (10 to 30 µg) and immunoprecipitated samples were combined with denaturing buffer and incubated at 100ºC for 10 min. Samples were digested *overnight* (*ON*) with 1 unit of PNGase F kit (New England BioLabs) at 37ºC, 300 rpm. The deglycosylated proteins were loaded onto 7,5% SDS-PAGE and immunoblotting with anti E-Cad or Con A lectin. In order to control the digestion process, samples were incubated in the same exact conditions but without the enzyme.

3.4 Western-Blot and lectin blot analysis

After reaching confluence status, cell cultures were submitted to a washing step with autoclaved Phosphate-Buffered Saline (PBS). Cell protein lysates were obtained by lysing cell cultures using a cold PBS solution containing 1%(v/v) Triton X-100, 1%(v/v) NP40, cOmplete (a protease inhibitor cocktail – Roche, 1 tablet/50 mL buffer) and a phosphatase inhibitor cocktail (composed by Phenylmethanesulfonyl fluoride (PMSF) and Vanadate - Sigma, 1:100 dilution). This lysis buffer was applied on the cell monolayers (that were kept on ice) for the period of 15 minutes. After that, the cells were scrapped and centrifuged at 13200 *rpm* for 10 minutes at 4ºC. Prior to use in a Western-Blot (WB) or lectin Blot experiment, total protein content of each lysate was quantified using a BCA protein assay kit (Pierce). Equal amounts of total cell protein lysates (10-35 µg) were submitted to 7,5% SDS-PAGE electrophoresis after the addition of Elution buffer (Laemmli 4X, β-mercaptoethanol and Bromophenol Blue) to each sample and performing a denaturing step (98ºC, 5 minutes). After electrophoresis, the separated proteins were transferred during 2 hours at 50V to nitrocellulose membranes (Amersham Biosciences). Ponceau solution was then used for rapid

staining of the membranes which were then immediately washed with a PBS-Tween 20 (PBST; Sigma Aldrich) 0,05% solution or with a Tris-Buffered Saline-Tween 20 (TBST) 0,05% solution.

The WB technique was performed using antibodies anti-E-Cad and anti-POMT2. After the washing of the Ponceau solution, membranes were blocked in a 5% milk solution in PBST 0,05% for E-Cad or in a 5% milk solution in TBST 0,05% for POMT2 for 1 hour. Afterwards, the mouse monoclonal anti-E-Cad antibody IgG2a (clone 36, BD Transduction Laboratories and Cell Signaling; 1:3000 dilution in 5% milk in PBST 0,05%) or the anti-POMT2 antibody (loop 5 of the protein, kindly provided by Sabine Strahl's group from University of Heidelberg; 1:500 dilution in 5% milk in TBST 0,05%) were used to incubate the membranes *ON* at 4ºC. At that time, the membranes were washed for 30 minutes (changing the solution every 10 minutes) at room temperature under agitation conditions. A PBST 0,05% solution was used for the E-Cad membranes, whereas in the case of the POMT2 membranes the washing steps were made with TBST 0,05% solution. For the E-Cad blot analysis, membranes were incubated for 1 hour at room temperature under agitation with horseradish peroxidase-linked goat anti-mouse IgG secondary antibody (Santa Cruz Biotechnology; 1:2000 dilution in 5% milk in PBST). For the POMT2 blot analysis, membranes were incubated for 1 hour at room temperature under agitation with horseradish peroxidase-linked goat anti-rabbit IgG secondary antibody (Santa Cruz Biotechnology; 1:2000 dilution in 5% milk in TBST). A final washing step, similar to what was previously described for each type of antibody used, was necessary before performing the detection of the immunoreactive bands using ECL reagent (Amersham Biosciences). Loading control analysis was done probing blots with anti-actin or anti-tubulin antibodies (Santa Cruz Biotechnology). The autoradiography films obtained from each independent experiment were analyzed in a densitometer. The bands of interest were quantified and normalized comparing to the control that was being used (actin or tubulin).

In order to perform the lectin blot analysis, membranes were blocked in a 4% bovine serum albumin (BSA; Sigma Aldrich) solution in PBST 0,05% for lectin blot *ON*. The lectin blots were then probed with Con A (Vector Laboratories) or L-PHA (Vector Laboratories) lectins for 1 hour at room temperature under agitation, and then were washed for 30 minutes (changing the solution every 10 minutes) at room temperature

under agitation conditions with a PBST 0,05% solution. While Con A recognizes α-Man residues present on the sample (either from *N*-glycans or *O*-Man glycans), L-PHA reacts to the presence of β-1,6 GlcNAc monosaccharides. To visualize the lectin immunoreactive bands, Vectorstain ABC kit (Vector Laboratories) was used to incubate the membranes for 1 hour at room temperature under agitation. The membranes were then washed 3 times (for 30 minutes) at room temperature under agitation conditions with a PBST 0,05% solution. The ABC kit is a solution of the complex avidin-peroxidase that binds with high affinity to the lectins we use as they are biotinylated, and biotin interacts with the avidin of the complex. The complexes antigen-lectin-biotin-avidin-peroxidase were then detected by adding ECL that works as a substrate for the peroxidase of the complex. Loading control analysis was done probing blots with anti-actin or anti-tubulin antibodies (Santa Cruz Biotechnology). The autoradiography films obtained from each independent experiment were analyzed in a densitometer.

3.5 Immunoprecipitation of E-Cadherin

For the IP protocols equal amounts of total protein from each cell lysate (750-1000 µg) were submitted to a *Preclear* step with 30 µL of protein G-sepharose beads (Sigma) for 1 hour under agitation. Afterwards, the samples were centrifuged (13200 *rpm*, 5 minutes, 4ºC) to obtain the supernatant and incubated *ON* with 1,5 µg of mouse monoclonal anti-E-Cad antibody IgG2a (clone 36, BD Transduction Laboratories and Cell Signaling) at 4ºC under agitation. Subsequently, the content of each eppendorf was incubated with 55 µL of pre-blocked protein G-sepharose beads (with a 1% BSA solution in IP buffer – 10%(v/v) lysis buffer and 1%(v/v) NaF 100X (10mM) in PBS) for 2 hours under agitation. After that, IP buffer was used to wash the beads-antibody-antigen complexes (3 times, 750 µL). The immune complexes (antibody-antigen) were then released by performing a boiling step (5 minutes, 98ºC) in Laemmli sampling buffer, and the final samples were subjected to a 7,5% SDS-PAGE electrophoresis. The separated proteins were then transferred to a nitrocellulose membrane, and the posthumous treatment of the membranes was similar to that of a normal WB or lectin blot procedure.

Each experiment was reproduced at least 2 times using protein extract from 2 completely different biological replicas. For all data comparisons, the Student's T-Test was used (two tailed, unequal variance).

3.6 Quantitative Real-Time PCR (qRT-PCR)

Total RNA from MKN45 MOCK and MKN45+GnT-V cells was extracted using Tri-Reagent (Sigma) according to the manufacturer's protocol. 1 µg of total RNA was reversed transcribed to single stranded cDNA using Superscript II Reverse Transcriptase and random hexamer primers (Invitrogen, Oregon, USA). Quantitative Real-Time-PCR (qRT-PCR) was carried out in triplicates for the target genes *POMT2* and *MGAT5*, and for the endogenous control *GAPDH* using as probe sets Hs00203575_m1, Hs00159136_m1 (Applied Biosystems, California, USA) and Hs.PT.58.40035104 (Integrated DNA Technologies, Iowa, USA) respectively. During data analysis, all *POMT2* and MGAT5 results were normalized against the control GAPDH. GaphPad Prism version 6.00 for Windows (GraphPad Software, La Jolla California USA, www.graphpad.com) was used to obtain the respective graphs.

4 RESULTS

4.1 MKN28 and KATO III display different morphologies and a different E-Cad patterns of expression

In order to acess the role of *O*-mannosylation in cancer, and particularly on E-Cad, we used two GC cell line models known to display different differentiation and phenotypical features, namely MKN28 and KATO III. MKN28 was first established from a moderately differentiated tubular adenocarcinoma from a 70-year old Asian female patient, while KATO III was derived from a poorly differentiated signet-ring cell carcinoma at a metastatic site from a 55-year old Asian male patient[102,103]. MKN28 and KATO III GC cell lines showed distinct morphologies and phenotypes when observed under bright field and after being cultured under the same conditions. MKN28 GC cells display an epithelial-like phenotype, forming a very well defined monolayer of cohesive cells that grow in adherence to the flask **(Figure 12A)**. Conversely, KATO III gastric carcinoma cells are characterized by being sparse, diffuse and small round-shaped, while some of those cells present a fibroblastoid-like appearance **(Figure 12B)**.

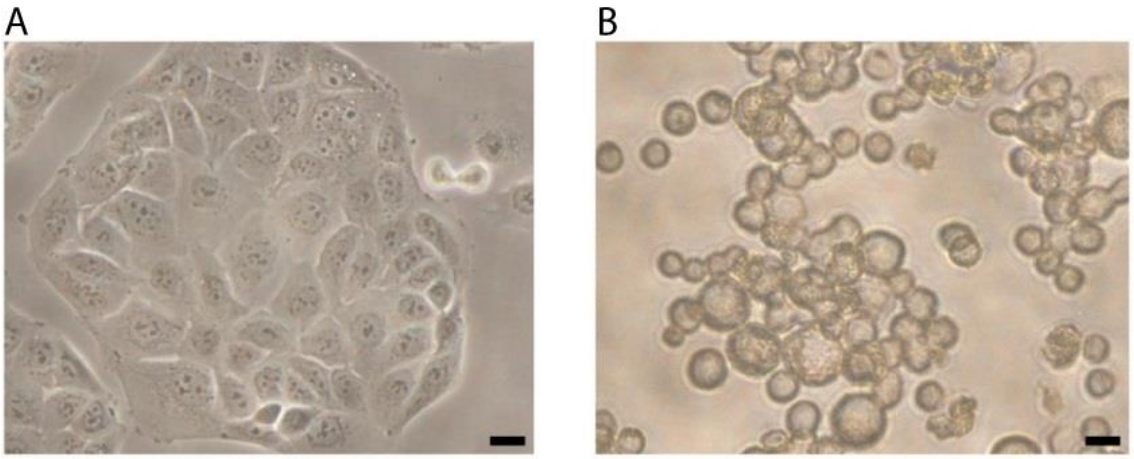

Figure 12 - MKN28 and KATO III cells morphological and phenotypical characteristics. (A) MKN28 cell line display an epithelial-like phenotype, forming a monolayer of adhesive cells while (B) KATO III cells are smaller and round-shaped.

The pattern of E-Cad expression in these two GC cell lines was assessed by E-Cad immunofluorescence. The results showed that MKN28 displayed a normal pattern of expression of E-Cad at the cell membrane (adhesion junctions) **(Figure 13A)**, whereas in KATO III we observed a mislocalization of E-Cad to the cytoplasm and we could clearly distinguish cells with the fibroblastoid-like phenotype **(Figure 13B)**, showing that E-Cad is normally expressed in MKN28 and clearly mislocalized in KATO III.

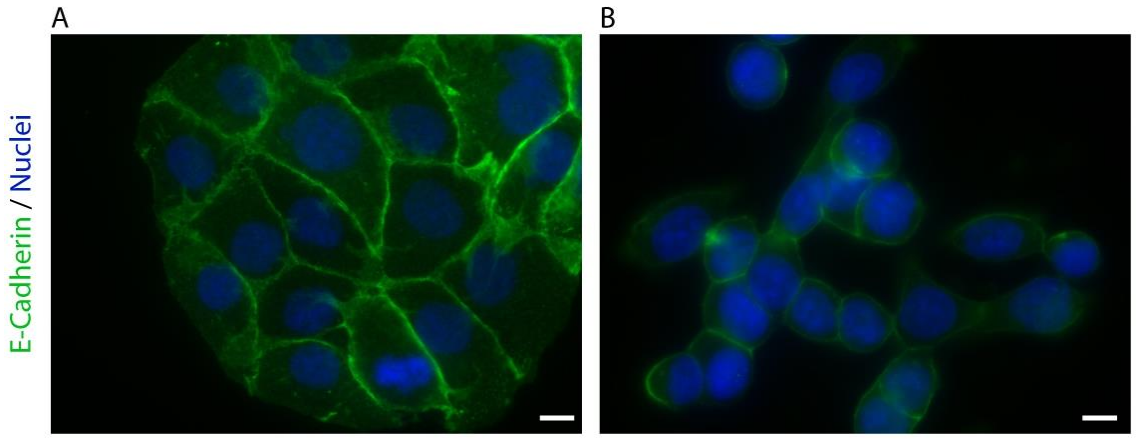

Figure 13 - Immunofluorescence analysis on (A) MKN28 and (B) KATO III cell lines. MKN28 cells demonstrated a cell membranous pattern of E-Cadherin expression, whereas KATO III cells displayed a mislocalization of the protein.

4.2 Impact of POMT2 protein expression levels variability in the overall *O*-mannosylation profile and in the E-Cadherin *O*-mannosylation status comparing MKN28 and KATO III cell line models

Taking into consideration the differential pattern of E-Cad expression in these two GC cell line models, together with previous studies[98,99] showing that E-Cad is a target of *O*-mannosylation, we went to access the *O*-mannosylation profile of E-Cad in these two GC contexts. We have evaluated the mRNA expression levels of POMT1 and POMT2 (data not shown) in those cell lines, and observed that POMT2's mRNA was expressed almost two times more in MKN28 comparing to KATO III, while POMT1's mRNA was expressed at similar levels. The levels of POMT2 protein expression were assessed using total protein lysates from MKN28 and KATO III cell lines. POMT2 is an ER transmembrane protein, formed by several helixes and loops. The antibody anti-POMT2 used targeted the loop5, the largest loop of this protein. The results showed that POMT2 is expressed in MKN28 cell line and that it is reduced in KATO III **(Figure 14)**.

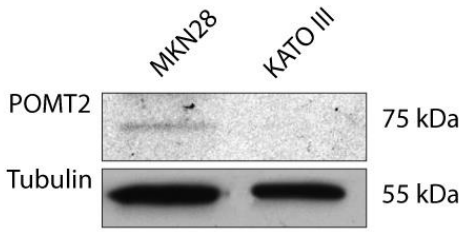

Figure 14 - Expression levels of POMT2 and the overall *O*-mannosylation profile. (a) Analysis by Western blot from total cell lysates of MKN28 and KATO III probed with antibodies against POMT2 and tubulin (loading control). POMT2 is expressed in MKN28 while in KATO III no visible band appears. (b) Lectin blot and Western blot analysis from digested total cell lysates of MKN28 and KATO III cells with Con A and the antibody against actin (loading control). MKN28 presents higher content in *O*-mannosyl glycans when compared to KATO III, as the lectin blot presents higher reactivity in the MKN28 cell line lane.

Afterwards, we have evaluated the overall *O*-mannosylation profile of these two different gastric cancer cell lines. In order to study the *O*-Man glycans accurately, a PNGase F digestion step was previously performed to remove all the *N*-linked glycans from all the proteins present in the total cell lysates, therefore increasing the specificity of Con A towards *O*-Man structures. Con A identifies all the α-Man residues present in the sample, which means that it does not discriminate between *O*- and *N*-linked α-Man residues, hence the necessity of trim all the *N*-glycans present on the sample. As observed in **Figure 15** the reactivity against Con A in the MKN28 cell line was completely distinct from the one obtained for KATO III, with MKN28 presenting higher content of proteins with *O*-mannosyl glycans attached, hence higher Con A reactivity.

The overall *O*-mannosylation status of the total cell lysates studied pointed to an increase in the number of *O*-mannosyl glycans present in the well differentiated MKN28 cell line, which led us to address the *O*-mannosylation profile specifically on E-Cad protein. In order to do so, we performed E-Cad immunoprecipitation (IP) of MKN28 and KATO III cell lines. Then, the E-Cad IP was digested with PNGase F to remove all the *N*-glycans possibly attached to E-Cad, and submitted the samples to a SDS-PAGE and Con A WB. We observed that for equal levels of E-Cad that was immunoprecipitated, the reactivity to Con A in the 120 kDa region of E-Cad from

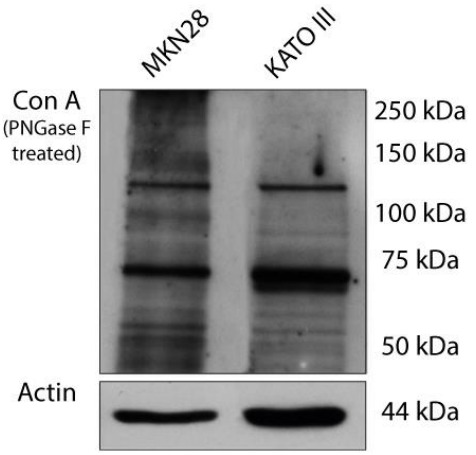

Figure 15 - The overall *O*-mannosylation profile of MKN28 and KATO III. Lectin blot and Western blot analysis from digested total cell lysates of MKN28 and KATO III cells with Con A and the antibody against actin (loading control). MKN28 presents higher content in *O*-mannosyl glycans when compared to KATO III, as the lectin blot presents higher reactivity in the MKN28 cell line lane.

MKN28 cell line was much higher (about three times more) when compared to the Con A reactivity in E-Cad from KATO III cell line **(Figure 16)**. Therefore, a GC cell line model

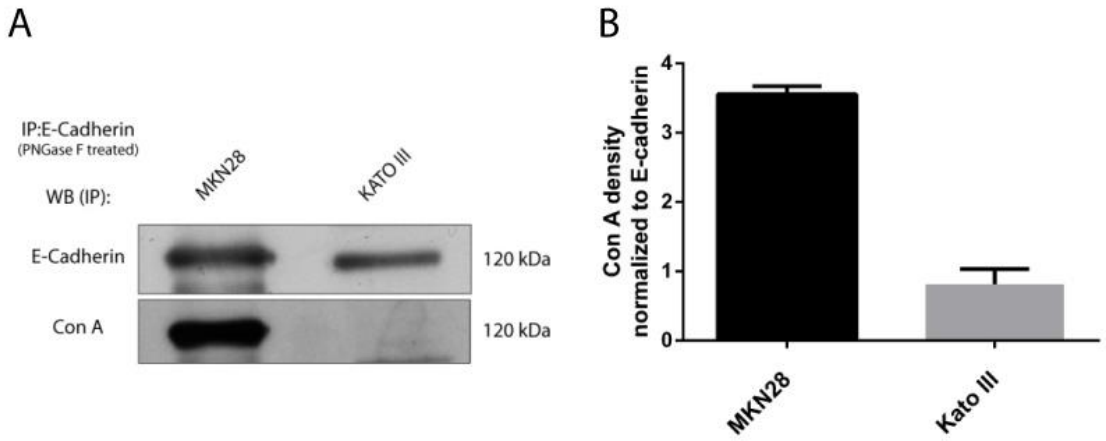

Figure 16 - *O*-mannosylation of E-Cadherin. (A) E-Cadherin was immunoprecipitated from MKN28 and KATO III cell lines and the existence of *O*-mannosyl glycans attached was assessed using simultaneously PNGase F digestion and Con A lectin blot. (B) *O*-mannosyl glycans were only detected on E-Cadherin in MKN28 cell line. Level of E-Cadherin was assessed in parallel using the antibody anti-E-Cadherin (clone 36) as a control of the immunoprecipitation. The levels of Con A density for MKN28 and KATO III were obtained after normalization to E-Cadherin.

that is moderately differentiated as MKN28 and that presents a normal pattern of E-Cad distribution at the cell membrane, has increased expression of POMT2 (both at the mRNA and the protein levels), as well as a high overall level of *O*-mannosylation of glycoproteins and increased *O*-Man presence on E-Cad. On the other hand, loss of differentiation of a cell line, such as in KATO III, is accompanied by low levels of POMT2 expression (both at the mRNA and the protein levels) and overall *O*-mannosylation, as well as a low levels of *O*-Man glycans on E-Cad.

4.3 Relationship between the levels of *O*-mannose glycans and β-1,6 GlcNAc branched *N*-glycans in different gastric cancer cell lines

Our group has demonstrated recently the importance of GnT-V activity in the impairment of E-Cad biological function in a GC context, namely due to the addition of β-1,6 GlcNAc branched *N*-glycans specifically on E-Cad-*N*-glycosylation site 1 (Asn554). Furthermore, it was also demonstrated that mutagenesis of this Asn by a glutamine

(Gln) residue and silencing of *MGAT5* have a protective effect, and prevents dysregulation of E-Cad by precluding the attachment of β-1,6 GlcNAc branching *N*-glycans at this site[27]. Taken together, these observations led us to explore the interplay between *N*-glycosylation and *O*-mannosylation on E-Cad protein in a GC context, especially due to the fact that several of those glycosylation sites (*N*-glycosylation and *O*-mannosylation) are close to each other[99].

We started by comparing the general profiles of *O*-mannosylation and *N*-glycosylation-mediated by GnT-V on both MKN28 and KATO III cell lines. For doing so, total cell lysates from these two cell lines were submitted to a lectin blotting using either Con A (after removal of *N*-glycans by PNGase F treatment) or L-PHA lectins, to evaluate the *O*-mannosyl glycans and β-1,6 GlcNAc branching *N*-glycans, respectively.

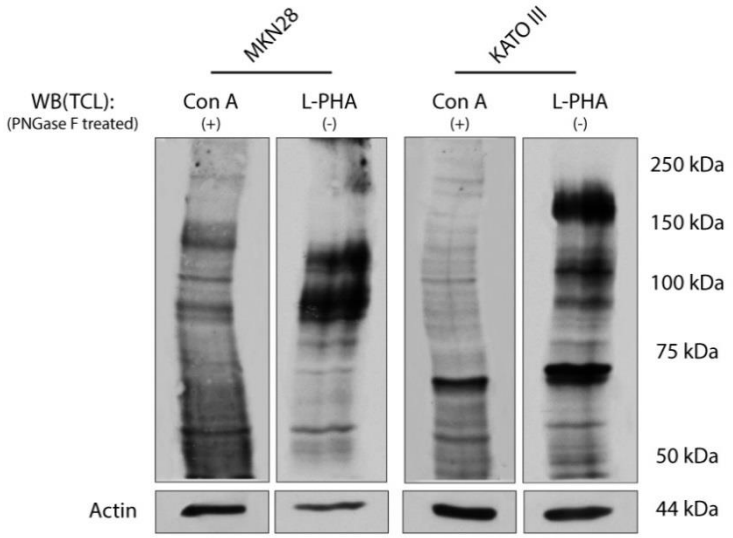

Figure 17 - Levels of *O*-mannosyl glycans and branched *N*-glycans in total cell lysates of MKN28 and KATO III cell lines. Lectin blots were performed in fresh total cell lysates, and actin was used as a loading control for both of them. An initial step of digestion was necessary when using Con A lectin to assess the *O*-mannosylation status of the lysates.

As depicted in **Figure 17**, the overall Con A reactivity in MKN28 is high comparing with L-PHA reactivity, while in KATO III we observed a reduced reactivity of Con A compairing with L-PHA. It is interesting that, considering the phenotype of each cell line, we have a higher level of *O*-Man glycans in the well-differentiated cells MKN28, while in the poorly-differentiated KATO III cells we have higher levels of β-1,6 GlcNAc branching *N*-glycans. Thus, these evidences seem to correlate with the observable phenotype for each cell line.

Afterwards we went to evaluate the levels of *O*-Man and β-1,6 GlcNAc branched structures specifically on E-Cad. We immunoprecipitated E-Cad using total cell lysates of MKN28 and KATO III cell lines, with subsequent digestion of the protein content that

was immunoprecipitated with PNGase F to remove the *N*-glycans that were attached to E-Cad, and submitted the samples to a WB protocol. We observed that, for equal levels of E-Cad that was immunoprecipitated, the level of Con A reactivity on MKN28 E-Cad band was higher than the levels of L-PHA reactivity. In the case of KATO III, we observed that Con A reactivity was decreased on E-Cad, while L-PHA reactivity on E-Cad band was positive **(Figure 18)**. These differential levels of *O*-Man and β-1,6 GlcNAc branching *N*-glycans on E-Cad seem to correlate with the differentiation/phenotype of these two GC cell lines and with E-Cad cellular localization. Moreover, it is interesting that, once again, we can correlate a more aggressive phenotype (KATO III) with loss of *O*-Man and the more differentiated phenotype (MKN28) with high levels of *O*-Man glycans on E-Cad.

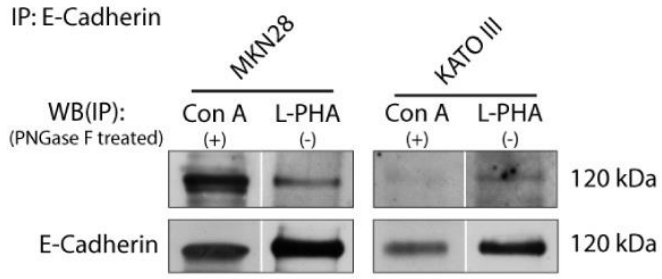

Figure 18 - E-Cadherin *O*-mannosylation and branched *N*-glycans content in MKN28 and KATO III cell lines. (A) E-Cadherin was immunoprecipitated in order to study its *O*-mannosylation profile and β-1,6 glycans content. Whereas MKN28 E-Cadherin presents a huge level of *O*-mannosyl glycans, it seems that in KATO III E-Cadherin is not very *O*-mannosylated. The levels of L-PHA reactivity in both cell lines are very similar to each other.

To further explore the relationship between β-1,6 GlcNAc branching *N*-glycans and *O*-mannosylation, we have used MKN45 MOCK and MKN45+GnT-V cell lines. MKN45 is a cell line that was derived in 1976 from a poorly differentiated adenocarcinoma of medullary type of a 62-year old female patient. This human gastric adenocarcinoma cell line was transfected in a stable way with the empty vector (MKN45 MOCK) or with GnT-V (MKN45+GnT-V) by Ihara *et al*. in 2002[101]. As GnT-V is the only difference between MKN45 MOCK and MKN45+GnT-V, and GnT-V is responsible for the production of β-1,6 GlcNAc branching *N*-glycans, this makes these two cell lines ideal models to study the influence of GnT-V on E-Cad. Here, we intend to use these two cell lines to clarify specifically if overexpression of GnT-V and high levels of β-1,6 GlcNAc branching *N*-glycans influence *O*-mannosylation of E-Cad.

In order to do so, a RT-PCR protocol was performed to study the expression of *MGAT5* and *POMT2* mRNA in MKN45 MOCK and MKN45+GnT-V cell lines. As the

results were normalized against an endogenous control, relative expression levels of *MGAT5* and *POMT2* mRNA could be compared in those cell lines. We confirmed a 30-fold increase in the expression level of *MGAT5* mRNA in the MKN45+GnT-V when compared to the MKN45 MOCK cell line **(Figure 19 *left*)**. When considering *POMT2* expression, we registered a slight increase (about 0,5 times more) in the relative expression of its mRNA in MKN45+GnT-V, when compared to MKN45 MOCK cell line **(Figure 19 *right*)**.

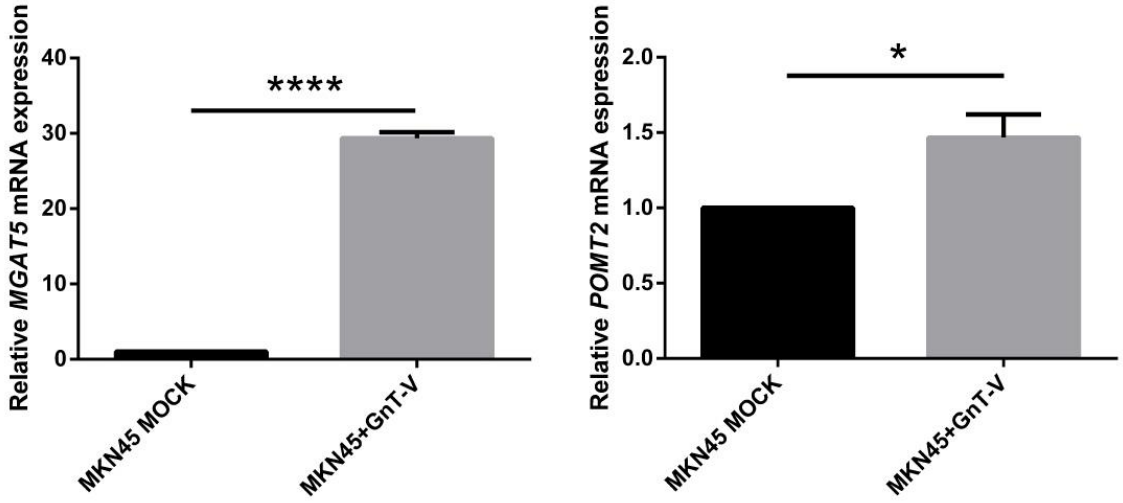

Figure 19 - Relative expression of *MGAT5* and *POMT2* mRNA in MKN45 MOCK and MKN45+GnT-V cell lines. (*left*) *MGAT5* is expressed approximately 30 times more in the MKN45 cell line transfected with GnT-V. (*right*) *POMT2* mRNA levels are slightly superior in MKN45+GnT-V when compared to the cells transfected with the empty vector (MKN45 MOCK). Results were normalized for *GAPDH* mRNA expression levels. Relative expression of MKN45 MOCK was considered to be 1, and the expression of MKN45+GnT-V was compared in each condition (Student's t-test: * P≤0.05; **** P≤0.001).

To continue exploring the interplay between β-1,6 GlcNAc branching *N*-glycosylation and *O*-mannosylation, we performed an IP of E-Cad using total cell lysates of these two cell lines, and ran in a 7,5% SDS-PAGE gel. The IP samples were used either to evaluate the *O*-mannosylation status of E-Cad (previously digested to remove the *N*-glycans content) with Con A, or to assess the presence of β-1,6 branched *N*-glycans on E-Cad with L-PHA **(Figure 20A)**. These results showed that MKN45+GnT-V cell line presents a significant and consistent decrease in *O*-mannosyl glycans attached to E-Cad when comparing to MKN45 MOCK cell line (almost to half of the structures) **(Figure 20B)**. At the same time, MKN45+GnT-V presents around two times more reactivity for L-PHA, which means higher levels of β-1,6 GlcNAc branched structured *N*-

glycans linked to E-Cad, in contrast to what happens in the MKN45 MOCK cell line **(Figure 20C)**.

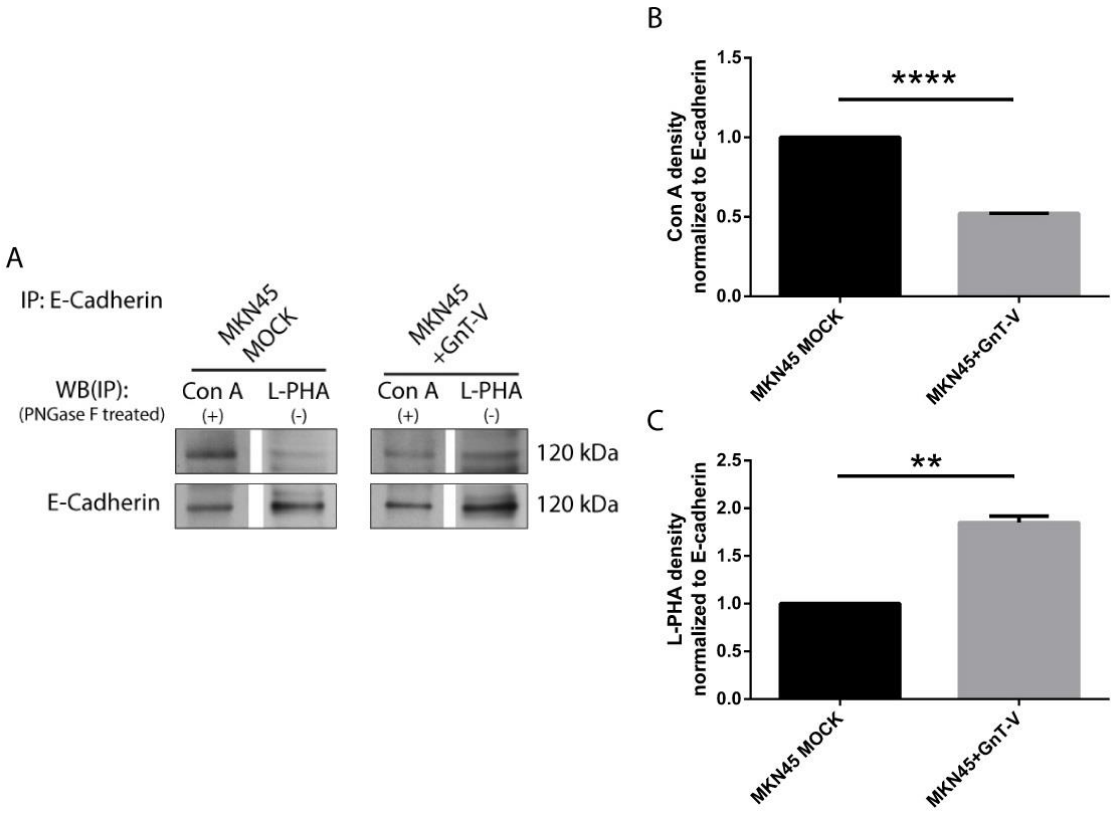

Figure 20 - E-Cadherin *O*-mannosylation and branched *N*-glycans content in MKN45 MOCK and MKN45+GnT-V cell lines. (A) E-Cadherin was immunoprecipitated in order to study its *O*-mannosylation profile and β-1,6 glycans content. The presence of GnT-V (MKN45+GnT-V) increases the number of β-1,6 branched *N*-glycans, when comparing to MKN45 MOCK. The level of reactivity of Con A is higher in MKN45 MOCK cells, which indicate a greater number of *O*-mannosyl glycans present on E-Cadherin, when compared to MKN45+GnT-V. (B and C) The levels of Con A and L-PHA densities for MKN45 MOCK and MKN45+GnT-V were obtained after normalization to E-Cadherin, being expressed as the fold change, compared with the association level of the same lectin with E-cadherin in MKN45 MOCK cells, which was taken as 1 (Student's t-test: **P≤0.01; **** P≤0.001).

5 DISCUSSION AND CONCLUSIONS

Discussion and Conclusions

The vast majority of cancer types present an aberrant glycosylation profile. In recent years, researchers have been paying more attention to the role of glycans in a cancer context due to the fact that many glycosyl epitopes are also tumor-associated antigens. The fact is glycobiology has been contributing to decipher important cancer mechanisms that may be applied to improve cancer diagnosis and in the development of new therapeutic strategies[51,72,104,105].

In a normal cellular environment, a set of glycosyltransferases regulates the synthesis of a group of specific glycans that can be attached to the proteins, whereas in a tumor microenvironment glycoproteins can be produced with a set of cancer-specific glycans, in a cell and tissue specific manner. For instance GnT-III and GnT-V act enzymatically on the same core structure linked to a protein [Asn]-GlcNAc2Man3$(GlcNAc)_2$[80]. However, while GnT-III is responsible for the addition of a bisecting GlcNAc and suppression of tumor metastasis[106], GnT-V adds the same monosaccharide but forming a β-1,6 GlcNAc branching structured glycan, described to have a huge importance in tumor progression and metastatic process[107,108]. It is also known that action of GnT-III always precedes that of GnT-V, and that the product that is formed cannot be then modified by a β-1,6 branching GlcNAc[81]. The expression levels of these glycosyltransferases have an impact on several cellular mechanisms, namely cell-cell adhesion[80,83,109].

The adherens junctions are essential to maintain tissue architecture and integrity through an array of intercellular signaling pathways[110,111]. E-Cad is one of the most important components of the adherens junctions, being composed by a transmembrane domain, a cytoplasmic domain, and a large extracellular domain that comprises the five cadherin repeats - ECs 1 to 5[21]. E-Cad is also a glycoprotein modified by *N*- and *O*-glycans that are important for its folding, expression and functions[25,26,98,99]. Dysfunction of E-Cad has, as expected, tremendous deleterious effects in cell biology, and has been considered as one of the key factors for cancer development and progression[42,43,112]. Throughout the years, our group was able to characterize in detail the role of alterations in the *N*-glycosylation profile of E-Cad, and how they influence the attainment of a malignant phenotype[27,40,51,78,108,113], backing up

the notion that the differential occupation of the four *N*-glycosylation sites available on E-Cad, individually or simultaneously, determines the modulation of E-Cad expression and its biological activity and function. Recently, *O*-mannosylation was related for the first time to cadherins, specifically E-Cad, by two individual laboratories: the Copenhagen Center for Glycomics discovered nine *O*-mannosylation sites available in E-Cad[98], while the role of *O*-Man on E-Cad biological function was reported at the Heidelberg University[99].

Taking these observations altogether, the present study intends to specifically define how the *O*-mannosylation profile of E-Cad has an impact on its biological function, in a cancer context. Moreover, the relation between *N*-glycosylation and *O*-mannosylation of E-Cad, in the same pathological context, was also explored.

We studied two different GC cell lines that display different morphological and phenotypical features, MKN28 and KATO III cell lines. MKN28 cell line was established from a moderately differentiated tubular adenocarcinoma that grows in a more epithelial-like fashion, while KATO III cell line is derived from a poorly differentiated advanced gastric carcinoma. Accordingly, the pattern of expression of E-Cad in each cell line model is distinguishable, with MKN28 displaying a well-defined membranar pattern of E-Cad expression and KATO III exhibiting a mislocalization of E-Cad **(Figures 12 and 13)**. Using these two different GC cell lines (well-differentiated, with normal E-Cad expression versus poorly-differentiated with mislocalization of E-Cad) as models we intend to explore the overall and E-Cad specific *O*-mannosylation profile of each cell line, and the interplay between *O*-mannosylation and *N*-glycosylation. We correlated both the POMT2 expression and the overall *O*-mannosylation profile of each cell line. The WB for POMT2 shows a strong relationship with the lectin blot for *O*-Man glycans, with MKN28 presenting higher expression of POMT2 and more reactivity for the Con A lectin when comparing to KATO III **(Figures 14 and 15)**. In fact, in KATO III a faint reactivity to POMT2 was detected using the antibody for the loop5 of the protein, which may potentially indicate its incorrect expression. Afterwards, we went to assess the pattern of *O*-Man on E-Cad in MKN28 and KATO III; so we immunoprecipitated the protein in order to proceed to an analysis of its *O*-mannosyl glycans content. PNGase F digestion of the immunoprecipitates was needed to trim all the *N*-glycans existent on E-Cad, and the lectin blot with Con A allowed us to detect specifically α-*O*-Man

monosaccharides. **Figure 16** reveals that E-Cad from MKN28 presents a high content in *O*-Man while poor *O*-mannosylation is detected on E-Cad from KATO III, which seems to indicate that E-Cad loses its *O*-Man content in a more aggressive phenotype.

Summing up the results presented so far, and taking into account the recognized importance of correct E-Cad *O*-mannosylation in the cell-cell adhesion process[99], we can correlate the expression of POMT2 with an higher level of *O*-mannosylated glycans in glycoproteins in general, but also in the specific case of E-Cad. These results seem to be in accordance to what was observed at Heidelberg University, as an increase in adherence of cells (MKN28 cell line) is accompanied by an enhancement in POMT2 expression and *O*-mannosylated glycoproteins. On the case of the more aggressive phenotype of KATO III, we can correlate it with the absence of POMT2 and of *O*-mannosylation of E-Cad.

The next step was to associate the *O*-mannosylation and the β-1,6 branched *N*-glycosylation levels in both GC cell line models, using a lectin blot strategy. We observed a distinguishable reactivity of Con A and L-PHA in MKN28 and KATO III, with the latter presenting a huge distribution of proteins with branched *N*-glycans and lower levels of *O*-mannosyl glycans. Conversely, MKN28 presents a high content of proteins that are *O*-mannosylated, and lower levels of β-1,6 GlcNAc branched *N*-glycans. We have also shown that MKN28 presents more *O*-Man content than KATO III, which made us realize that the overall distribution of *O*-Man and β-1,6 *N*-glycans can be inversely correlated, and that the expression of higher contents of *O*-Man also corresponds to a more epithelial-like phenotype **(Figure 17)**. IP of E-Cad would provide us the perfect assessment of the overall glycosylation profile of the protein. In several experiments, we have evaluated the *O*-Man distribution and the β-1,6 branched *N*-glycans presence on E-Cad individually. The *O*-Man content was almost absent in E-Cad from KATO III lysates, but there is a high reactivity towards L-PHA, indicating the presence of a higher content of β-1,6 branching structures on E-Cad. On the other hand, E-Cad from MKN28 exhibits a huge reactivity towards Con A, pointing to the presence of several *O*-Man glycans, while L-PHA reactivity is lower, indicating the presence of less β-1,6 GlcNAc branching structures on E-Cad **(Figure 18)**.

To further validate these results relating the relationship between *O*-mannosylation and *N*-glycosylation, we have used two other cancer cell line models, MKN45 MOCK

and MKN45+GnT-V, as they have already been widely used by our group as important models for studying the influence of β-1,6 GlcNAc glycans on E-Cad in GC[108]. To confirm the overexpression of GnT-V in MKN45+GnT-V, we performed a RT-PCR protocol for the *MGAT5* mRNA of MKN45 MOCK and MKN45+GnT-V cell lines. As previously demonstrated[108], the latter presents an almost 30-fold increase in mRNA relative expression when compared to the MOCK cells **(Figure 19 – left)**. We also analyzed *POMT2* mRNA expression using the same technique, and concluded that there is a slight increase in the MKN45+GnT-V cell line **(Figure 19 – right)**. Afterwards, we immunoprecipitated E-Cad from both cell lines to evaluate specifically the *O*-Man and β-1,6 branched *N*-glycans content. Considering MKN45 MOCK we have a higher content of *O*-Man residues attached to E-Cad, and a smaller reactivity to L-PHA that indicates reduced amounts of β-1,6 branched *N*-glycans linked to E-Cad in this cell line when compared to MKN45+GnT-V **(Figure 20)**. In light to what was described by our group the results suggest that an increase in *O*-Man content and the absence of branched *N*-glycans attached to E-Cad leads to a more stable phenotype[108]. However, the role of GnT-III in this conundrum remains to be clarified, as it would be interesting to assess if we would get an increase in *O*-mannosylation of E-Cad by the insertion of GnT-III or, on the other hand, the occupation of the *N*-glycosylation sites by bisecting *N*-glycans could lead to a halt in the *O*-mannosylation process.

Overall, it seems that POMT2 protein has in fact a role in E-Cad *O*-mannosylation in cancer and in the modulation of its functions[67,114]. Interestingly, the cancer cell line with a more aggressive phenotype (KATO III) loses *O*-Man glycans attached to E-Cad, which is associated with its cellular mislocalization and in line to what was presented by Sabine Strahl's group regarding the importance of *O*-mannosylation of E-Cad for cell-cell adhesion[99]. The connection between *O*-mannosylation and *N*-glycosylation is also an innovative finding particularly in a GC context. The presence of *O*-Man on E-Cad appears to be inversely related with the presence of β-1,6 branched *N*-glycans, and the aggressiveness of a cancer might be determined by the balance of these two types of glycosylation **(Figure 21)**.

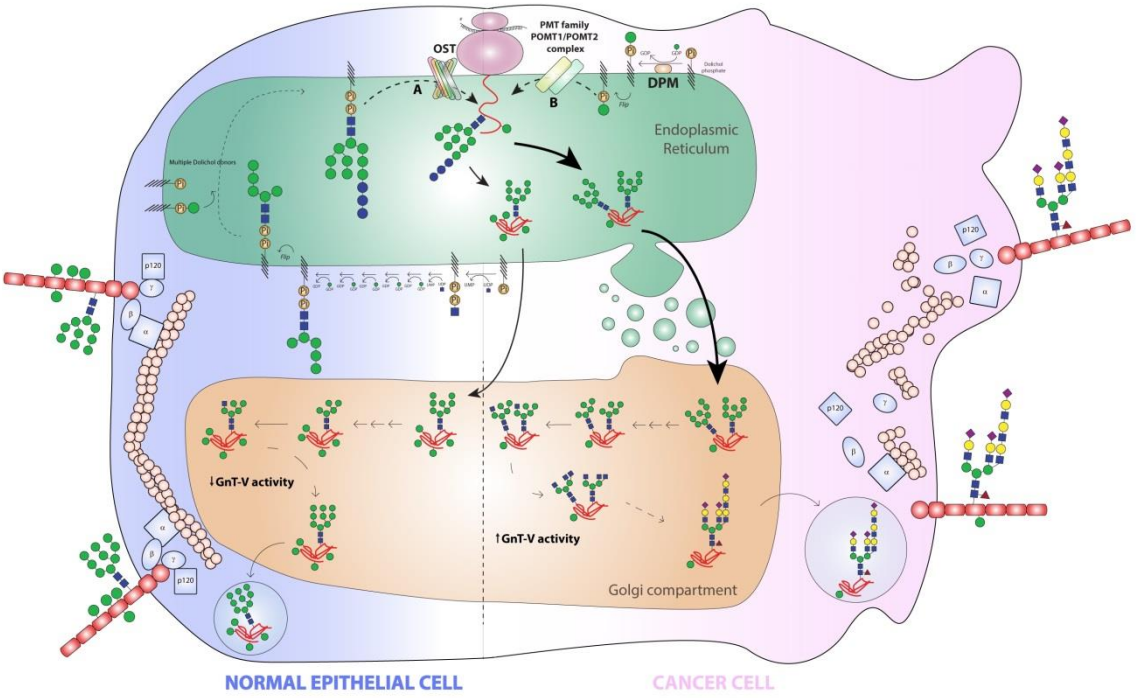

Figure 21 - Proposed model for the interplay between *O*-mannosylation and *N*-glycosylation. In a normal epithelial cell context, more *O*-mannosyl glycans are linked to E-Cad, as the protein does not have such bulky *N*-glycans due to the absence of GnT-V activity. On the other hand, in an adenocarcinoma cell context, GnT-V activity leads to the formation of huge complex *N*-glycan structures, which precludes the attachment of so many *O*-Man residues, contributing to a decrease in cell-cell adhesion and to a decreased stability of the E-Cad-catenins complex.

[These results are a part of a more comprehensive project. Part of these results was included in a manuscript that will be submitted for publication.]

6 FUTURE PERSPECTIVES

Future perspectives

Glycobiology plays a huge role in cancer, as further corroborated by this study. The prominence of *N*-glycosylation in tumor cell-cell adhesion, specifically the importance of modifications of E-Cad *N*-glycans structure, has already been extensively studied[12,27,78,108]. The novelty surrounding *O*-mannosylation allows researchers to hypothesize new ways of how to address old problems. During the course of this project, we intended to assess an unexplored scientific issue on the role of *O*-mannosylation in GC, focusing on the modulation of the functions of E-Cad. Indeed, we were able to establish a new field of expertise relating glycobiology and cancer, as no one has ever described *O*-mannosylation mechanisms in this pathological context. Moreover, E-Cad *O*-mannosylation profile, and the interplay with *N*-glycosylation, was also studied in different cancer contexts. Some of the mechanisms behind this relationship however remain unraveled.

When the Copenhagen group proposed that E-Cad could be strongly *O*-mannosylated, they did it using a SimpleCell technique that allowed them to look solely to *O*-Man monosaccharides present on glycoproteins without further extension[98]. However, in a cellular environment we do not know what actually happens in terms of *O*-mannosylation proteins such as E-Cad. In this case, it is of the utmost importance to determine the preferential sites of *O*-mannosylation of E-Cad, because maybe between the 9 identified positions and the 6 that were predicted we can have dominance in the locale of *O*-mannosylation, as verified for *N*-glycosylation[27]. The fact is, *O*-mannosyl glycans can exist in those positions, but, in a cell that composes an organ we do not know how many of these glycans we have, where, and how they modify the E-Cad biological function.

It is also important to know how *O*-mannosylation and *N*-glycosylation of E-Cad are specifically connected, as several of these sites are close to one another as depicted in **Figure 10**. Does, for instance, any *N*-glycan (high-mannose content, hybrid or complex) that occupies Asn554 affect the attachment of *O*-Man to Thr505? We have proven here that the introduction of a glycosyltransferase (GnT-V) that produces branching *N*-glycans in a cell line actually induces a decrease in the number of *O*-Man monosaccharides on E-Cad. But what is the mechanism behind this? Could it be just

the occupancy of the *N*-glycosylation sites, or to a regulation of the levels of POMT2 protein or other glycosyltransferases produced? It is also very important to explore the machinery that regulates POMT2 expression (as well as other POMTs) and if it is in fact this protein the sole responsible for the *O*-mannosylation of E-Cad.

While studying, we have actually started to paying more interest to POMT2 when compared to POMT1, as we performed an RT-PCR to evaluate the mRNA levels of *POMT1* and *POMT2* genes and realized that, when comparing MKN28 and KATO III cell lines, *POMT2* mRNA was more expressed in MKN28, while *POMT1* showed no significant differences amongst the two cell lines (results not shown). We chose to specifically learn about the role of POMT2 protein on *O*-mannosylation in GC, but we obtained some intriguing results **(Figure 15)** that make us question the relative importance of each mannosyltransferase to the overall process, and if there are other redundancies unknown, such as other POMTs not yet described.

Due to the huge importance of E-Cad on a GC context, all the knowledge relating its glycans content has enormous impact, especially if we can find a distinguishable pattern between a normal situation and a disease state, envisioning clinical applications.

7 REFERENCES

References

1 Torre, L. A. *et al.* Global cancer statistics, 2012. *CA: a cancer journal for clinicians* **65**, 87-108, doi:10.3322/caac.21262 (2015).

2 Ferlay, J. *et al.* (International Agency for Research on Cancer, Lyon, France, 2015).

3 Pinho, S. S. *et al.* Gastric cancer: adding glycosylation to the equation. *Trends in molecular medicine* **19**, 664-676, doi:10.1016/j.molmed.2013.07.003 (2013).

4 Lauren, P. The Two Histological Main Types of Gastric Carcinoma: Diffuse and So-Called Intestinal-Type Carcinoma. An Attempt at a Histo-Clinical Classification. *Acta pathologica et microbiologica Scandinavica* **64**, 31-49 (1965).

5 Vauhkonen, M., Vauhkonen, H. & Sipponen, P. Pathology and molecular biology of gastric cancer. *Best practice & research. Clinical gastroenterology* **20**, 651-674, doi:10.1016/j.bpg.2006.03.016 (2006).

6 Hu, B. *et al.* Gastric cancer: Classification, histology and application of molecular pathology. *Journal of gastrointestinal oncology* **3**, 251-261, doi:10.3978/j.issn.2078-6891.2012.021 (2012).

7 Cancer Genome Atlas Research, N. Comprehensive molecular characterization of gastric adenocarcinoma. *Nature* **513**, 202-209, doi:10.1038/nature13480 (2014).

8 Correa, P. Human gastric carcinogenesis: a multistep and multifactorial process--First American Cancer Society Award Lecture on Cancer Epidemiology and Prevention. *Cancer research* **52**, 6735-6740 (1992).

9 Anastas, J. N. & Moon, R. T. WNT signalling pathways as therapeutic targets in cancer. *Nature Reviews Cancer* **13**, 11-26, doi:Doi 10.1038/Nrc3419 (2013).

10 Carneiro, F. *et al.* Model of the early development of diffuse gastric cancer in E-cadherin mutation carriers and its implications for patient screening. *The Journal of pathology* **203**, 681-687, doi:10.1002/path.1564 (2004).

11 Oliveira, C., Seruca, R. & Carneiro, F. Genetics, pathology, and clinics of familial gastric cancer. *International journal of surgical pathology* **14**, 21-33 (2006).

12 Paredes, J. *et al.* Epithelial E- and P-cadherins: role and clinical significance in cancer. *Biochimica et biophysica acta* **1826**, 297-311, doi:10.1016/j.bbcan.2012.05.002 (2012).

13 Wadhwa, R. *et al.* Gastric cancer-molecular and clinical dimensions. *Nature reviews. Clinical oncology* **10**, 643-655, doi:10.1038/nrclinonc.2013.170 (2013).

14 Duraes, C., Almeida, G. M., Seruca, R., Oliveira, C. & Carneiro, F. Biomarkers for gastric cancer: prognostic, predictive or targets of therapy? *Virchows Archiv : an international journal of pathology* **464**, 367-378, doi:10.1007/s00428-013-1533-y (2014).

15 Takeichi, M. Functional correlation between cell adhesive properties and some cell surface proteins. *The Journal of cell biology* **75**, 464-474 (1977).

16 Hyafil, F., Morello, D., Babinet, C. & Jacob, F. A cell surface glycoprotein involved in the compaction of embryonal carcinoma cells and cleavage stage embryos. *Cell* **21**, 927-934 (1980).

17 Hyafil, F., Babinet, C. & Jacob, F. Cell-cell interactions in early embryogenesis: a molecular approach to the role of calcium. *Cell* **26**, 447-454 (1981).

18 Yoshida-Noro, C., Suzuki, N. & Takeichi, M. Molecular nature of the calcium-dependent cell-cell adhesion system in mouse teratocarcinoma and embryonic cells studied with a monoclonal antibody. *Developmental biology* **101**, 19-27 (1984).

19 Hatta, K. & Takeichi, M. Expression of N-cadherin adhesion molecules associated with early morphogenetic events in chick development. *Nature* **320**, 447-449, doi:10.1038/320447a0 (1986).

20 Berx, G. *et al.* Cloning and characterization of the human invasion suppressor gene E-cadherin (CDH1). *Genomics* **26**, 281-289 (1995).

21 van Roy, F. & Berx, G. The cell-cell adhesion molecule E-cadherin. *Cellular and molecular life sciences : CMLS* **65**, 3756-3788, doi:10.1007/s00018-008-8281-1 (2008).

22 Nagafuchi, A., Shirayoshi, Y., Okazaki, K., Yasuda, K. & Takeichi, M. Transformation of cell adhesion properties by exogenously introduced E-cadherin cDNA. *Nature* **329**, 341-343, doi:10.1038/329341a0 (1987).

23 Shapiro, L. *et al.* Structural basis of cell-cell adhesion by cadherins. *Nature* **374**, 327-337, doi:10.1038/374327a0 (1995).

24 Ozawa, M., Hoschutzky, H., Herrenknecht, K. & Kemler, R. A possible new adhesive site in the cell-adhesion molecule uvomorulin. *Mechanisms of development* **33**, 49-56 (1990).

25 Pinho, S. S. *et al.* Modulation of E-cadherin function and dysfunction by N-glycosylation. *Cellular and molecular life sciences : CMLS* **68**, 1011-1020, doi:10.1007/s00018-010-0595-0 (2011).

26 Zhou, F. *et al.* Unglycosylation at Asn-633 made extracellular domain of E-cadherin folded incorrectly and arrested in endoplasmic reticulum, then sequentially degraded by ERAD. *Glycoconjugate journal* **25**, 727-740, doi:10.1007/s10719-008-9133-9 (2008).

27 Carvalho, S. *et al.* Preventing E-cadherin aberrant N-glycosylation at Asn-554 improves its critical function in gastric cancer. *Oncogene*, doi:10.1038/onc.2015.225 (2015).

28 Schackmann, R. C., Tenhagen, M., van de Ven, R. A. & Derksen, P. W. p120-catenin in cancer - mechanisms, models and opportunities for intervention. *Journal of cell science* **126**, 3515-3525, doi:10.1242/jcs.134411 (2013).

29 Gumbiner, B. M. Regulation of cadherin-mediated adhesion in morphogenesis. *Nature reviews. Molecular cell biology* **6**, 622-634, doi:10.1038/nrm1699 (2005).

30 Guillot, C. & Lecuit, T. Mechanics of epithelial tissue homeostasis and morphogenesis. *Science* **340**, 1185-1189, doi:10.1126/science.1235249 (2013).

31 Becker, K. F. *et al.* Exon skipping in the E-cadherin gene transcript in metastatic human gastric carcinomas. *Human molecular genetics* **2**, 803-804 (1993).

32 Takeichi, M. Cadherins in cancer: implications for invasion and metastasis. *Current opinion in cell biology* **5**, 806-811 (1993).

33 Becker, K. F. *et al.* E-cadherin gene mutations provide clues to diffuse type gastric carcinomas. *Cancer research* **54**, 3845-3852 (1994).

34 Oda, T. *et al.* E-cadherin gene mutations in human gastric carcinoma cell lines. *Proceedings of the National Academy of Sciences of the United States of America* **91**, 1858-1862 (1994).

35 Tamura, G. *et al.* Inactivation of the E-cadherin gene in primary gastric carcinomas and gastric carcinoma cell lines. *Japanese journal of cancer research : Gann* **87**, 1153-1159 (1996).

36 Handschuh, G. *et al.* Tumour-associated E-cadherin mutations alter cellular morphology, decrease cellular adhesion and increase cellular motility. *Oncogene* **18**, 4301-4312, doi:10.1038/sj.onc.1202790 (1999).

37 Ascano, J. J. *et al.* Inactivation of the E-cadherin gene in sporadic diffuse-type gastric cancer. *Modern pathology : an official journal of the United States and Canadian Academy of Pathology, Inc* **14**, 942-949, doi:10.1038/modpathol.3880416 (2001).

38 Yoshiura, K. *et al.* Silencing of the E-cadherin invasion-suppressor gene by CpG methylation in human carcinomas. *Proceedings of the National Academy of Sciences of the United States of America* **92**, 7416-7419 (1995).

39 Korpal, M., Lee, E. S., Hu, G. & Kang, Y. The miR-200 family inhibits epithelial-mesenchymal transition and cancer cell migration by direct targeting of E-cadherin transcriptional repressors ZEB1 and ZEB2. *The Journal of biological chemistry* **283**, 14910-14914, doi:10.1074/jbc.C800074200 (2008).

40 Pinho, S. S. *et al.* Role of E-cadherin N-glycosylation profile in a mammary tumor model. *Biochemical and biophysical research communications* **379**, 1091-1096, doi:10.1016/j.bbrc.2009.01.024 (2009).

41 Carneiro, P. *et al.* E-cadherin dysfunction in gastric cancer--cellular consequences, clinical applications and open questions. *FEBS letters* **586**, 2981-2989, doi:10.1016/j.febslet.2012.07.045 (2012).

42 Hanahan, D. & Weinberg, R. A. Hallmarks of cancer: the next generation. *Cell* **144**, 646-674, doi:10.1016/j.cell.2011.02.013 (2011).

43 Christofori, G. New signals from the invasive front. *Nature* **441**, 444-450, doi:10.1038/nature04872 (2006).

44 Fujita, Y. *et al.* Hakai, a c-Cbl-like protein, ubiquitinates and induces endocytosis of the E-cadherin complex. *Nature cell biology* **4**, 222-231, doi:10.1038/ncb758 (2002).

45 Hartsock, A. & Nelson, W. J. Competitive regulation of E-cadherin juxtamembrane domain degradation by p120-catenin binding and Hakai-mediated ubiquitination. *PloS one* **7**, e37476, doi:10.1371/journal.pone.0037476 (2012).

46 Joo, Y. E. *et al.* Expression of e-cadherin and catenins in early gastric cancer. *Journal of clinical gastroenterology* **35**, 35-42 (2002).

47 Corso, G. *et al.* Somatic mutations and deletions of the E-cadherin gene predict poor survival of patients with gastric cancer. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology* **31**, 868-875, doi:10.1200/JCO.2012.44.4612 (2013).

48 Varki, A. Biological roles of oligosaccharides: all of the theories are correct. *Glycobiology* **3**, 97-130 (1993).

49 Varki, A. *et al.* in *Essentials of Glycobiology* (eds A. Varki *et al.*) (Cold Spring Harbor Laboratory Press, 2009).

50 Moremen, K. W., Tiemeyer, M. & Nairn, A. V. Vertebrate protein glycosylation: diversity, synthesis and function. *Nature reviews. Molecular cell biology* **13**, 448-462, doi:10.1038/nrm3383 (2012).

51 Pinho, S. S. & Reis, C. A. Glycosylation in cancer: mechanisms and clinical implications. *Nature reviews. Cancer*, doi:10.1038/nrc3982 (2015).

52 Gloster, T. M. Advances in understanding glycosyltransferases from a structural perspective. *Current opinion in structural biology* **28**, 131-141, doi:10.1016/j.sbi.2014.08.012 (2014).

53 Kornfeld, R. & Kornfeld, S. Assembly of asparagine-linked oligosaccharides. *Annual review of biochemistry* **54**, 631-664, doi:10.1146/annurev.bi.54.070185.003215 (1985).

54 Aebi, M., Bernasconi, R., Clerc, S. & Molinari, M. N-glycan structures: recognition and processing in the ER. *Trends in biochemical sciences* **35**, 74-82, doi:10.1016/j.tibs.2009.10.001 (2010).

55 Hart, G. W. Dynamic O-linked glycosylation of nuclear and cytoskeletal proteins. *Annual review of biochemistry* **66**, 315-335, doi:10.1146/annurev.biochem.66.1.315 (1997).

56 Endo, T. O-mannosyl glycans in mammals. *Biochimica et biophysica acta* **1473**, 237-246 (1999).

57 Helenius, A. & Aebi, M. Intracellular functions of N-linked glycans. *Science* **291**, 2364-2369 (2001).

58 Helenius, A. & Aebi, M. Roles of N-linked glycans in the endoplasmic reticulum. *Annual review of biochemistry* **73**, 1019-1049, doi:10.1146/annurev.biochem.73.011303.073752 (2004).

59 Parodi, A. J. Role of N-oligosaccharide endoplasmic reticulum processing reactions in glycoprotein folding and degradation. *The Biochemical journal* **348 Pt 1**, 1-13 (2000).

60 Clausen, H. & Bennett, E. P. A family of UDP-GalNAc: polypeptide N-acetylgalactosaminyl-transferases control the initiation of mucin-type O-linked glycosylation. *Glycobiology* **6**, 635-646 (1996).

61 Bennett, E. P. *et al.* Control of mucin-type O-glycosylation: a classification of the polypeptide GalNAc-transferase gene family. *Glycobiology* **22**, 736-756, doi:10.1093/glycob/cwr182 (2012).

62 Gill, D. J., Clausen, H. & Bard, F. Location, location, location: new insights into O-GalNAc protein glycosylation. *Trends in cell biology* **21**, 149-158, doi:10.1016/j.tcb.2010.11.004 (2011).

63 Loibl, M. *et al.* Protein O-mannosyltransferases associate with the translocon to modify translocating polypeptide chains. *The Journal of biological chemistry* **289**, 8599-8611, doi:10.1074/jbc.M113.543116 (2014).

64 Manya, H. *et al.* Demonstration of mammalian protein O-mannosyltransferase activity: coexpression of POMT1 and POMT2 required for enzymatic activity. *Proceedings of the National Academy of Sciences of the United States of America* **101**, 500-505, doi:10.1073/pnas.0307228101 (2004).

65 Takahashi, S. *et al.* A new beta-1,2-N-acetylglucosaminyltransferase that may play a role in the biosynthesis of mammalian O-mannosyl glycans. *Glycobiology* **11**, 37-45 (2001).

66 Yoshida-Moriguchi, T. *et al.* SGK196 is a glycosylation-specific O-mannose kinase required for dystroglycan function. *Science* **341**, 896-899, doi:10.1126/science.1239951 (2013).

67 Praissman, J. L. & Wells, L. Mammalian O-mannosylation pathway: glycan structures, enzymes, and protein substrates. *Biochemistry* **53**, 3066-3078, doi:10.1021/bi500153y (2014).

68 Bleckmann, C. *et al.* O-glycosylation pattern of CD24 from mouse brain. *Biological chemistry* **390**, 627-645, doi:10.1515/BC.2009.044 (2009).

69 Stalnaker, S. H. *et al.* Glycomic analyses of mouse models of congenital muscular dystrophy. *The Journal of biological chemistry* **286**, 21180-21190, doi:10.1074/jbc.M110.203281 (2011).

70 Lee, J. K. *et al.* Developmental expression of the neuron-specific N-acetylglucosaminyltransferase Vb (GnT-Vb/IX) and identification of its in vivo glycan products in comparison with those of its paralog, GnT-V. *The Journal of biological chemistry* **287**, 28526-28536, doi:10.1074/jbc.M112.367565 (2012).

71 Ohtsubo, K. & Marth, J. D. Glycosylation in cellular mechanisms of health and disease. *Cell* **126**, 855-867, doi:10.1016/j.cell.2006.08.019 (2006).

72 Reis, C. A., Osorio, H., Silva, L., Gomes, C. & David, L. Alterations in glycosylation as biomarkers for cancer detection. *Journal of clinical pathology* **63**, 322-329, doi:10.1136/jcp.2009.071035 (2010).

73 Taniguchi, N. & Korekane, H. Branched N-glycans and their implications for cell adhesion, signaling and clinical applications for cancer biomarkers and in therapeutics. *BMB reports* **44**, 772-781 (2011).

74 Adamczyk, B., Tharmalingam, T. & Rudd, P. M. Glycans as cancer biomarkers. *Biochimica et biophysica acta* **1820**, 1347-1353, doi:10.1016/j.bbagen.2011.12.001 (2012).

75 Yoshimura, M., Ihara, Y., Matsuzawa, Y. & Taniguchi, N. Aberrant glycosylation of E-cadherin enhances cell-cell binding to suppress metastasis. *The Journal of biological chemistry* **271**, 13811-13815 (1996).

76 Liwosz, A., Lei, T. & Kukuruzinska, M. A. N-glycosylation affects the molecular organization and stability of E-cadherin junctions. *The Journal of biological chemistry* **281**, 23138-23149, doi:10.1074/jbc.M512621200 (2006).

77 Akama, R. *et al.* N-acetylglucosaminyltransferase III expression is regulated by cell-cell adhesion via the E-cadherin-catenin-actin complex. *Proteomics* **8**, 3221-3228, doi:10.1002/pmic.200800038 (2008).

78 Pinho, S. S. *et al.* The role of N-acetylglucosaminyltransferase III and V in the post-transcriptional modifications of E-cadherin. *Human molecular genetics* **18**, 2599-2608, doi:10.1093/hmg/ddp194 (2009).

79 Zhao, H. *et al.* N-glycosylation affects the adhesive function of E-Cadherin through modifying the composition of adherens junctions (AJs) in human breast carcinoma cell line MDA-MB-435. *Journal of cellular biochemistry* **104**, 162-175, doi:10.1002/jcb.21608 (2008).

80 Gu, J. *et al.* A mutual regulation between cell-cell adhesion and N-glycosylation: implication of the bisecting GlcNAc for biological functions. *Journal of proteome research* **8**, 431-435, doi:10.1021/pr800674g (2009).

81 Zhao, Y. *et al.* N-acetylglucosaminyltransferase III antagonizes the effect of N-acetylglucosaminyltransferase V on alpha3beta1 integrin-mediated cell migration. *The Journal of biological chemistry* **281**, 32122-32130, doi:10.1074/jbc.M607274200 (2006).

82 Zhao, Y. Y. *et al.* Functional roles of N-glycans in cell signaling and cell adhesion in cancer. *Cancer science* **99**, 1304-1310, doi:10.1111/j.1349-7006.2008.00839.x (2008).

83 Taniguchi, N., Miyoshi, E., Ko, J. H., Ikeda, Y. & Ihara, Y. Implication of N-acetylglucosaminyltransferases III and V in cancer: gene regulation and signaling mechanism. *Biochimica et biophysica acta* **1455**, 287-300 (1999).

84 Granovsky, M. *et al.* Suppression of tumor growth and metastasis in Mgat5-deficient mice. *Nature medicine* **6**, 306-312, doi:10.1038/73163 (2000).

85 Iijima, J. *et al.* Cell-cell interaction-dependent regulation of N-acetylglucosaminyltransferase III and the bisected N-glycans in GE11 epithelial cells. Involvement of E-cadherin-mediated cell adhesion. *The Journal of biological chemistry* **281**, 13038-13046, doi:10.1074/jbc.M601961200 (2006).

86 de-Freitas-Junior, J. C. *et al.* Insulin/IGF-I signaling pathways enhances tumor cell invasion through bisecting GlcNAc N-glycans modulation. an interplay with E-cadherin. *PloS one* **8**, e81579, doi:10.1371/journal.pone.0081579 (2013).

87 Guo, H. B., Johnson, H., Randolph, M. & Pierce, M. Regulation of homotypic cell-cell adhesion by branched N-glycosylation of N-cadherin extracellular EC2 and EC3 domains. *The Journal of biological chemistry* **284**, 34986-34997, doi:10.1074/jbc.M109.060806 (2009).

88 Stowell, S. R., Ju, T. & Cummings, R. D. Protein glycosylation in cancer. *Annual review of pathology* **10**, 473-510, doi:10.1146/annurev-pathol-012414-040438 (2015).

89 Sentandreu, R. & Northcote, D. H. The structure of a glycopeptide isolated from the yeast cell wall. *The Biochemical journal* **109**, 419-432 (1968).

90 Strahl-Bolsinger, S., Immervoll, T., Deutzmann, R. & Tanner, W. PMT1, the gene for a key enzyme of protein O-glycosylation in Saccharomyces cerevisiae. *Proceedings of the National Academy of Sciences of the United States of America* **90**, 8164-8168 (1993).

91 Gentzsch, M. & Tanner, W. Protein-O-glycosylation in yeast: protein-specific mannosyltransferases. *Glycobiology* **7**, 481-486 (1997).

92 Martin-Blanco, E. & Garcia-Bellido, A. Mutations in the rotated abdomen locus affect muscle development and reveal an intrinsic asymmetry in Drosophila. *Proceedings of the National Academy of Sciences of the United States of America* **93**, 6048-6052 (1996).

93 Strahl-Bolsinger, S., Gentzsch, M. & Tanner, W. Protein O-mannosylation. *Biochimica et biophysica acta* **1426**, 297-307 (1999).

94 Jurado, L. A., Coloma, A. & Cruces, J. Identification of a human homolog of the Drosophila rotated abdomen gene (POMT1) encoding a putative protein O-mannosyl-transferase, and assignment to human chromosome 9q34.1. *Genomics* **58**, 171-180, doi:10.1006/geno.1999.5819 (1999).

95 Endo, T. Structure, function and pathology of O-mannosyl glycans. *Glycoconjugate journal* **21**, 3-7, doi:10.1023/B:GLYC.0000043740.26062.2c (2004).

96 Wells, L. The o-mannosylation pathway: glycosyltransferases and proteins implicated in congenital muscular dystrophy. *The Journal of biological chemistry* **288**, 6930-6935, doi:10.1074/jbc.R112.438978 (2013).

97 Beltran-Valero de Bernabe, D. *et al.* Mutations in the O-mannosyltransferase gene POMT1 give rise to the severe neuronal migration disorder Walker-Warburg syndrome. *American journal of human genetics* **71**, 1033-1043, doi:10.1086/342975 (2002).

98 Vester-Christensen, M. B. *et al.* Mining the O-mannose glycoproteome reveals cadherins as major O-mannosylated glycoproteins. *Proceedings of the National Academy of Sciences of the United States of America* **110**, 21018-21023, doi:10.1073/pnas.1313446110 (2013).

99 Lommel, M. *et al.* Protein O-mannosylation is crucial for E-cadherin-mediated cell adhesion. *Proceedings of the National Academy of Sciences of the United States of America* **110**, 21024-21029, doi:10.1073/pnas.1316753110 (2013).

100 Willer, T. *et al.* Targeted disruption of the Walker-Warburg syndrome gene Pomt1 in mouse results in embryonic lethality. *Proceedings of the National Academy of Sciences of the United States of America* **101**, 14126-14131, doi:10.1073/pnas.0405899101 (2004).

101 Ihara, S. *et al.* Prometastatic effect of N-acetylglucosaminyltransferase V is due to modification and stabilization of active matriptase by adding beta 1-6 GlcNAc branching. *The Journal of biological chemistry* **277**, 16960-16967, doi:10.1074/jbc.M200673200 (2002).

102 Sekiguchi, M., Sakakibara, K. & Fujii, G. Establishment of cultured cell lines derived from a human gastric carcinoma. *The Japanese journal of experimental medicine* **48**, 61-68 (1978).

103 Motoyama, T., Hojo, H. & Watanabe, H. Comparison of seven cell lines derived from human gastric carcinomas. *Acta pathologica japonica* **36**, 65-83 (1986).

104 Hakomori, S. Glycosylation defining cancer malignancy: new wine in an old bottle. *Proceedings of the National Academy of Sciences of the United States of America* **99**, 10231-10233, doi:10.1073/pnas.172380699 (2002).

105 Fuster, M. M. & Esko, J. D. The sweet and sour of cancer: glycans as novel therapeutic targets. *Nature reviews. Cancer* **5**, 526-542, doi:10.1038/nrc1649 (2005).

106 Yoshimura, M., Nishikawa, A., Ihara, Y., Taniguchi, S. & Taniguchi, N. Suppression of lung metastasis of B16 mouse melanoma by N-acetylglucosaminyltransferase III gene transfection. *Proceedings of the National Academy of Sciences of the United States of America* **92**, 8754-8758 (1995).

107 Dennis, J. W., Laferte, S., Waghorne, C., Breitman, M. L. & Kerbel, R. S. Beta 1-6 branching of Asn-linked oligosaccharides is directly associated with metastasis. *Science* **236**, 582-585 (1987).

108 Pinho, S. S. *et al.* E-cadherin and adherens-junctions stability in gastric carcinoma: functional implications of glycosyltransferases involving N-glycan branching biosynthesis, N-acetylglucosaminyltransferases III and V. *Biochimica et biophysica acta* **1830**, 2690-2700, doi:10.1016/j.bbagen.2012.10.021 (2013).

109 Guo, H. B., Lee, I., Kamar, M. & Pierce, M. N-acetylglucosaminyltransferase V expression levels regulate cadherin-associated homotypic cell-cell adhesion and intracellular signaling pathways. *The Journal of biological chemistry* **278**, 52412-52424, doi:10.1074/jbc.M308837200 (2003).

110 Harris, T. J. An introduction to adherens junctions: from molecular mechanisms to tissue development and disease. *Sub-cellular biochemistry* **60**, 1-5, doi:10.1007/978-94-007-4186-7_1 (2012).

111 Buckley, C. D. *et al.* Cell adhesion. The minimal cadherin-catenin complex binds to actin filaments under force. *Science* **346**, 1254211, doi:10.1126/science.1254211 (2014).

112 Christofori, G. Changing neighbours, changing behaviour: cell adhesion molecule-mediated signalling during tumour progression. *The EMBO journal* **22**, 2318-2323, doi:10.1093/emboj/cdg228 (2003).

113 Pinho, S. S. *et al.* Loss and recovery of Mgat3 and GnT-III Mediated E-cadherin N-glycosylation is a mechanism involved in epithelial-mesenchymal-epithelial transitions. *PloS one* **7**, e33191, doi:10.1371/journal.pone.0033191 (2012).

114 Loibl, M. & Strahl, S. Protein O-mannosylation: what we have learned from baker's yeast. *Biochimica et biophysica acta* **1833**, 2438-2446, doi:10.1016/j.bbamcr.2013.02.008 (2013).